Lecture Notes in Networks and Systems 693

Xin-She Yang
R. Simon Sherratt
Nilanjan Dey
Amit Joshi *Editors*

Proceedings of Eighth International Congress on Information and Communication Technology

ICICT 2023, London, Volume 1

Springer

Lecture Notes in Networks and Systems

Volume 693

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Xin-She Yang · R. Simon Sherratt · Nilanjan Dey · Amit Joshi
Editors

Proceedings of Eighth International Congress on Information and Communication Technology

ICICT 2023, London, Volume 1

Springer

Editors
Xin-She Yang
Department of Design Engineering
and Mathematics
Middlesex University London
London, UK

Nilanjan Dey
Department of Computer Science
and Engineering
Techno International New Town
Chakpachuria, West Bengal, India

R. Simon Sherratt
Department of Biomedical Engineering
University of Reading
Reading, UK

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-99-3242-9 ISBN 978-981-99-3243-6 (eBook)
https://doi.org/10.1007/978-981-99-3243-6

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The Eighth International Congress on Information and Communication Technology will be held during 20–23 February 2023, in a hybrid mode, physical at London, UK and digital platform: Zoom. ICICT 2023 was organised by Global Knowledge Research Foundation and managed by G. R. Scholastic LLP. The associated partners were Springer and InterYIT IFIP. The conference will provide a useful and wide platform both for display of the latest research and for exchange of research results and thoughts. The participants of the conference will be from almost every part of the world, with backgrounds of either academia or industry, allowing a real multinational and multicultural exchange of experiences and ideas.

A great pool of more than 1300 papers were received for this conference from across 113 countries among which around 361 papers were accepted and will be presented physically at London and digital platform Zoom during the four days. Due to the overwhelming response, we had to drop many papers in the hierarchy of the quality. Total 46 technical sessions will be organised in parallel in four days along with a few keynotes and panel discussions in hybrid mode. The conference will be involved in deep discussion and issues which will be intended to solve at global levels. New technologies will be proposed, experiences will be shared, and future solutions for design infrastructure for ICT will also be discussed. The final papers will be published in four volumes of proceedings by Springer LNNS Series. Over the years, this congress has been organised and conceptualised with collective efforts of a large number of individuals. I would like to thank each of the committee members and the reviewers for their excellent work in reviewing the papers. Grateful acknowledgements are extended to the team of Global Knowledge Research Foundation for their valuable efforts and support.

I look forward to welcoming you to the eighth edition of this ICICT Congress 2023.

Amit Joshi, Ph.D.
Organising Secretary, ICICT 2023
Director—Global Knowledge Research Foundation
Ahmedabad, India

Contents

Editors and Contributors

About the Editors

Xin-She Yang obtained his D.Phil. in Applied Mathematics from the University of Oxford. He then worked at Cambridge University and National Physical Laboratory (UK) as a Senior Research Scientist. Now he is Reader at Middlesex University London, Fellow of the Institute of Mathematics and its Application (IMA), and a Book Series co-editor of the *Springer Tracts in Nature-Inspired Computing*. He was also the IEEE Computational Intelligence Society task force chair for Business Intelligence and Knowledge Management (2015–2020). He has published more than 25 books and more than 400 peer-reviewed research publications with over 78,600 citations, and he has been on the prestigious list of highly-cited researchers (Web of Sciences) for seven consecutive years (2016–2022).

R. Simon Sherratt was born near Liverpool, England, in 1969. He is currently a Professor of Biosensors at the Department of Biomedical Engineering, University of Reading, UK. His main research area is signal processing and personal communications in consumer devices, focusing on wearable devices and health care. Professor Sherratt received the 1st place IEEE Chester Sall Memorial Award in 2006, the 2nd place in 2016 and the 3rd place in 2017.

Nilanjan Dey is an Associate Professor at the Department of Computer Science and Engineering, Techno International New Town, India. He is the Editor-in-Chief of the *International Journal of Ambient Computing and Intelligence*; a Series co-editor of *Springer Tracts in Nature-Inspired Computing* (STNIC), *Data-Intensive Research* (DIR), Springer Nature; and a Series co-editor of *Advances in Ubiquitous Sensing Applications for Healthcare*, Elsevier. He is a fellow of IETE and a Senior Member of IEEE.

Amit Joshi is currently the Director of Global Knowledge Research Foundation, also an Entrepreneur and Researcher who has completed his Masters and research in the areas of cloud computing and cryptography in medical imaging. Dr. Joshi has an experience of around 10 years in academic and industry in prestigious organisations. Dr. Joshi is an active member of ACM, IEEE, CSI, AMIE, IACSIT-Singapore, IDES, ACEEE, NPA and many other professional societies. Currently, Dr. Joshi is the International Chair of InterYIT at International Federation of Information Processing (IFIP, Austria), He has presented and published more than 50 papers in national and international journals/conferences of IEEE and ACM. Dr. Joshi has also edited more than 40 books which are published by Springer, ACM and other reputed publishers. Dr. Joshi has also organised more than 50 national and international conferences and programs in association with ACM, Springer, IEEE to name a few across different countries including India, UK, Europe, USA, Canada, Thailand, Egypt and many more.

Contributors

Sa'ed Abed Department of Computer Engineering, College of Engineering and Petroleum, Kuwait University, Kuwait City, Kuwait

Tesneem Abou-Kassem United Arab Emirates University, Al Ain, UAE

Stavolo Agostino University of Naples "Federico II", Naples, Italy

Mohammed Aissaoui National School of Applied Sciences, Mohammed Premier University, Oujda, Morocco

Muhmmad Al-Khiza'ay Al-Kitab University, Kerkuk-Altun Kupri, Iraq

Asmaa Abdul-Razzaq Al-Qaisi College of Education for Women, Baghdad University, Baghdad, Iraq

Noora Alallaq Al-Kitab University, Kerkuk-Altun Kupri, Iraq

Md. Golam Rabiul Alam BRAC University, Dhaka, Bangladesh

Jaflah Alammary University of Bahrain, Sakhir, Kingdom of Bahrain

Fatima Hamad Obaid Alazeezi United Arab Emirates University, Al Ain, UAE

Mohammed Alhayani Al-Noor University College, Ninevah-Mosul, Iraq

Mohamad Yusoff Alias Multimedia University, Cyberjaya, Malaysia

Sascha Alpers FZI Forschungszentrum Informatik, Karlsruhe, Germany

Kirti Ameta Department of Computer Science and Information Technology, JRN Rajasthan Vidyapeeth (Deemed to Be) University, Udaipur, Rajasthan, India

Paula Amorim Faculty of Health Sciences, University of Beira Interior, Covilhã, Portugal;
Rehabilitation Medicine Center of Central Region, Monte Redondo, Portugal

Victor H. Andaluz Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Tiago Martins Andrade BT Applied Research, Adastral Park, UK

Gohad Atul Anil Bosch Global Software Technologies, Bangalore, India

Alexandre F. J. Antunes Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança (IPB), Campus de Santa Apolónia, Bragança, Portugal

K. R. Anupama Department of Electrical and Electronics Engineering, BITS Pilani, Zuarinagar, Goa, India

Norizan Anwar School of Information Science, College of Computing, Informatics and Media, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

Ashrf Aoad Istanbul Sabahattin Zaim University, Istanbul, Turkey

Oluwasegun Julius Aroba ICT and Society Research Group, Information Systems Department, Durban University of Technology, Durban, South Africa;
Honorary Research Associate, Department of Operations and Quality Management, Faculty of Management Science, Durban University of Technology, Durban, South Africa

Adeleh Asemi Department of Software Engineering, Faculty of Computer Science and Information Technology, Universiti Malaya, Kuala Lumpur, Malaysia

Asefeh Asemi Doctoral School of Economics, Business, and Informatics, Corvinus University of Budapest, Budapest, Hungary

Diego Avila-Pesantez Grupo de Investigación en Innovación Científica y Tecnológica (GIICYT), Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador

Anwar Awang Man Faculty of Management, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia

Nahed Azab The American University in Cairo, New Cairo, Egypt

Rihards Balodis Institute of Mathematics and Computer Science, University of Latvia, Riga, Latvia

David P. Barros-Piedra Universidad Politécnica Salesiana, Cuenca, Ecuador

Yaser Baseri University of New Brunswick, Fredericton, Canada

Daniel Berleant University of Arkansas at Little Rock, Little Rock, AR, USA

Jeanne Bernaldo School of Information Technology, Mapúa University, Manila, Philippines

Sylvain Bernard Sandia National Laboratories, Albuquerque, NM, USA

Om Prakash Bharti Department of Electrical Engineering, Government Polytechnic College, Ghazipur, UP, India

Aditi Bhatia-Kalluri Faculty of Information, University of Toronto, Toronto, ON, Canada

Jonathan Bisila Sandia National Laboratories, Albuquerque, NM, USA

Eric Blancaflor School of Information Technology, Mapúa University, Manila, Philippines

Hajar Bnouachir Engineering Research Laboratory (LRI), System Architecture Team (EAS) National and High School of Electricity and Mechanic (ENSEM), Hassan II University, Research Foundation for Development and Innovation in Science and Engineering, Casablanca, Morocco

Ahmed Bouajaj Laboratory of Engineering Sciences and Applications, National School of Applied Sciences, Abdelmalek Essaâdi University, Al-Hoceima, Morocco

Michael Robert Buche Sandia National Laboratories, Albuquerque, NM, USA

Ndumiso Buthelezi Audit and Taxation, Audit and Account Management Department, Durban University of Technology, Durban, South Africa

Elijah Lowell Calip School of Information Technology, Mapúa University, Manila, Philippines

Kordian Caplazi Rimon Technologies GmbH, Zurich, Switzerland

Sonia Cárdenas-Delgado Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Yousra Chabchoub ISEP—Institut Supérieur d'Électronique de Paris, Paris, France

Ajay Kumar Chadayan Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Fa-Hsiang Chang Wenzhou-Kean University, Wenzhou, China

Hanna Charankevich University of Virginia, Arlington, VA, USA

Paul A. Chasi-Pesántez Universidad Politécnica Salesiana, Cuenca, Ecuador

Ab Razak Che Hussin Faculty of Management, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia

Mohammed Amine Chenouf National School of Applied Sciences, Mohammed Premier University, Oujda, Morocco

Meriyem Chergui Computer Science and Smart Systems (C3S), National and High School of Electricity and Mechanic (ENSEM) Hassan II University, Casablanca, Morocco

Chekwas Ifeanyi Chikezie Department of Telecommunications Engineering, Federal University of Technology, Minna, Niger State, Nigeria

Collence Takaingenhamo Chisita ICT and Society Research Group, Information Systems Department, Durban University of Technology, Durban, South Africa

Asif Hasan Chowdhury BRAC University, Dhaka, Bangladesh

Kwang Sik Chung Department of Computer Science, Korea National Open University, Seoul, Korea

Edwin J. Coronel-González Universidad Politécnica Salesiana, Cuenca, Ecuador

Diana Dacova Technical University of Sofia, Sofia, Bulgaria

A'Qilah Ahmad Dahalan CONFIRM Centre for SMART Manufacturing, University of Limerick, Limerick, Ireland;
Centre for Defence Foundation Studies, Universiti Pertahanan Nasional Malaysia, Kuala Lumpur, Malaysia

Natalia Danilina Department of Therapy and Endocrinology, RUDN University, Moscow, Russia

Michael David Department of Telecommunications Engineering, Federal University of Technology, Minna, Niger State, Nigeria

Ahmet Demirelli Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul, Türkiye

Leonel D. Deusdado Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança (IPB), Campus de Santa Apolónia, Bragança, Portugal

Gilhwan Do C&P Korea Co. Ltd, Busan, Republic of Korea

Ervin Domazet International Balkan University, Skopje, Macedonia

Menachem Domb Ashkelon Academy College, Ashkelon, Israel

Bouchra El Asri IMS Team, ADMIR Laboratory, Rabat IT Center, ENSIAS, Mohammed V University, Rabat, Morocco

Hussam Elbehiery Vanridge University, Denver, CO, USA

Khaled Elbehiery Devry University, Denver, CO, USA

Mohamed ElSherif Extend The Ad Network, New Maadi, Egypt

Su-Hong Eom Department of Electronic Engineering, Tech University of Korea, Siheung, Republic of Korea

Gurdal Ertek College of Business and Economics, United Arab Emirates University, Al Ain, UAE

Sandeep Kumar Erudiyanathan Bosch Global Software Technologies, Bangalore, India

Fairuz Faculty of Medicine and Health Sciences, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Szilard Zsolt Fazekas Department of Mathematical Science and Electrical-Electronic-Computer Engineering, Akita University, Akita, Japan

Chariguamán Quinteros Magali Fernanda Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Yusri Mahbub Firdaus Faculty of Industrial Technology Institut Teknologi Bandung (ITB), Bandung, Indonesia

Fadia Fitriyanti Master of Law, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Colin Flanagan University of Limerick, Limerick, Ireland

J. Flórez Universidad de los Andes, Bogotá, Colombia

Leonor Portugal da Fonseca Department of Informatics Engineering, University of Coimbra, Centre for Informatics and Systems of the University of Coimbra, Coimbra, Portugal

Grassia Maria Gabriella University of Naples “Federico II”, Naples, Italy

Andrea Gallardo Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Jason M. Gates Sandia National Laboratories, Albuquerque, NM, USA

Loay E. George University of Information Technology and Communication, Baghdad, Iraq

Mohammed Ghanim University of Bahrain, Sakhir, Kingdom of Bahrain

Joy Gisler ETH Zurich, Zurich, Switzerland

Joshua Goldstein University of Virginia, Arlington, VA, USA

Samuel Andrew Grayson Sandia National Laboratories, Albuquerque, NM, USA

Luis F. Guerrero-Vásquez Universidad Politécnica Salesiana, Cuenca, Ecuador

Shirin Gulova Department of Therapy and Endocrinology, RUDN University, Moscow, Russia

Arnav Gupta Symbiosis International (Deemed) University, Pune, India

Aritra Halder Drexel University, Philadelphia, PA, USA

Evan Harvey Sandia National Laboratories, Albuquerque, NM, USA

Osman Hasan School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan

Faiyaz Bin Hashem BRAC University, Dhaka, Bangladesh

Mohamed Khalafalla Hassan Innovation research and development center (IRDC), The Future University, Khartoum, Sudan

Kazumi Hatayama Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Christian Hirt ETH Zurich, Zurich, Switzerland

Lars Holmberg Department of Computer Science and Media Technology, Malmö University, Malmö, Sweden

Valentin Holzwarth RhySearch, Buchs, Switzerland

Mutaz Hamad Hussein Innovation research and development center (IRDC), The Future University, Khartoum, Sudan

İlter İrdesel Magneti Marelli, Bursa, Türkiye

Muhammad Amien Ibrahim Bina Nusantara University, Jakarta, Indonesia

Tamotsu Ichikawa ROHM Semiconductor, Yokohama, Japan

Muchammad Ichsan Master of Law, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Daisuke Iimori Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Niddal Imam Alfaisal University, Riyadh, Saudi Arabia

Takashi Ishida ROHM Semiconductor, Yokohama, Japan

Md. Fahim Islam BRAC University, Dhaka, Bangladesh

Alexander Jaeger Sandia National Laboratories, Albuquerque, NM, USA

Chasen Jeffries Claremont Graduate University, Claremont, CA, USA

Tang Mui Joo Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur, Malaysia

Sujata Joshi Symbiosis International (Deemed) University, Pune, India

Juhari Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Lakshmi Kailas College of Business and Economics, United Arab Emirates University, Al Ain, UAE

Srinivasan Kandhasamy Bosch Global Software Technologies, Bangalore, India

Mohammed E. A. Kanona Innovation research and development center (IRDC), The Future University, Khartoum, Sudan

Jada El Kasri Laboratory of Applied Geophysics, Geotechnics, Engineering and Environmental Geology (L3GIE), The Mohammadia School of Engineers, Mohammed V University, Rabat, Morocco

Shogo Katayama Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Kentaroh Katoh Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Barjinder Kaur University of New Brunswick, Fredericton, Canada

Alaoui Belghiti Khaoula Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

Ammar Kheirbek ISEP—Institut Supérieur d'Électronique de Paris, Paris, France

Dae-We Kim Department of Electronic Engineering, Tech University of Korea, Siheung, Republic of Korea

Ga-Young Kim Department of Electronic Engineering, Tech University of Korea, Siheung, Republic of Korea

Soeun Kim Sookmyung Women's University, Seoul, Republic of Korea

Hiroki Kimura Tokai University, Hiratsuka, Kanagawa, Japan

Andrea Ko Corvinus University of Budapest, Budapest, Hungary

Haruo Kobayashi Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Frantisek Kobzik Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung, Germany

Elizabeth Koumpan IBM Consulting, Markham, ON, Canada

Karina Kowarsch Claremont Graduate University, Claremont, CA, USA

Robert Kudelić Faculty of Organization and Information Science, Varaždin, Croatia

Takahiro Kunii Kashina System Co, Hikone, Japan

Andreas Kunz ETH Zurich, Zurich, Switzerland

Stefan Kunze Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung, Germany

Rangith B. Kuriakose Central University of Technology, Bloemfontein, Free State, South Africa

Cahyadi Kurniawan Government Science, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Yohannes Kurniawan Information Systems Department, School of Information Systems, Bina Nusantara University, Jakarta, Indonesia

Irina Kurnikova Department of Therapy and Endocrinology, RUDN University, Moscow, Russia;
Department of Aviation and Space Medicine, Federal State Budgetary Educational Institution of Further Professional Education, Russian Medical Academy of Continuous Professional Education, Moscow, Russia

Kusbaryanto Master of Hospital Adminstration, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Anna Kuwana Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Abdelaziz Lahmili Laboratory of Applied Geophysics, Geotechnics, Engineering and Environmental Geology (L3GIE), The Mohammadia School of Engineers, Mohammed V University, Rabat, Morocco

Kirk Timothy Landin Sandia National Laboratories, Albuquerque, NM, USA

Penio Dimitrov Lebamovski Institute of Robotics, Bulgarian Academy of Sciences, Sofia, Bulgaria

Jongwoo Lee Sookmyung Women's University, Seoul, Republic of Korea

Joo-Hyung Lee Department of Computer Engineering, Tech University of Korea, Siheung, Republic of Korea

Kwangil Lee Korea Maritime and Ocean University, Busan, Republic of Korea

Won-Young Lee Department of Electronic Engineering, Tech University of Korea, Siheung, Republic of Korea

Ahmet Lekesiz Faculty of Engineering, Marmara University, Istanbul, Türkiye

Sergei Levchenko International Institute of Applied Research and Technology, Sindelfingen, Germany

Meiyu Li Southeast University, Nanjing, China

Jinling Liang Southeast University, Nanjing, China

Xinyuan Lin Wenzhou-Kean University, Wenzhou, China

Xiaoxiao Liu University of Limerick, Limerick, Ireland

Mauricio Loachamín-Valencia Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Júlio C. Lopes Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança (IPB), Campus de Santa Apolónia, Bragança, Portugal

Rui P. Lopes Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança (IPB), Campus de Santa Apolónia, Bragança, Portugal

C. Lozano-Garzón Universidad de los Andes, Bogotá, Colombia

Yiyang Luo V. N. Karazin Kharkiv National University, Kharkiv, Ukraine

Irina Lutsenko O.Ya. Usikov Institute for Radiophysics and Electronics of the National Academy of Sciences of Ukraine, Kharkiv, Ukraine

Vladislav Lutsenko O.Ya. Usikov Institute for Radiophysics and Electronics of the National Academy of Sciences of Ukraine, Kharkiv, Ukraine

Chikkamath Manjunath Bosch Global Software Technologies, Bangalore, India

Marino Marina University of Naples "Federico II", Naples, Italy

Rhanoui Maryem Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

Mohamad Noorman Masrek School of Information Science, College of Computing, Informatics and Media, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

Praveen Mathew Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Jumpei Matsuura Nara Gakuen University, Nara City, Nara Prefecture, Japan

Jeffrey McCann Dell Technologies, Limerick, Ireland

Sean McGrath University of Limerick, Limerick, Ireland

Daniela Mechkaroska University of Information Science and Technology "St. Paul the Apostle", Ohrid, Macedonia

Hicham Medromi Engineering Research Laboratory (LRI), System Architecture Team (EAS) National and High School of Electricity and Mechanic (ENSEM), Hassan II University, Research Foundation for Development and Innovation in Science and Engineering, Casablanca, Morocco

Lora Metanova The St. Kliment Ohridski Sofia University, Sofia, Bulgaria

Mounia Mikram Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

Reed Milewicz Sandia National Laboratories, Albuquerque, NM, USA

Nadezhda Miteva The St. Kliment Ohridski Sofia University, Sofia, Bulgaria

S. M. Mizanur Rahman Bangladesh University of Professionals, Dhaka, Bangladesh

Khalid Sheikhidris Mohamed Innovation research and development center (IRDC), The Future University, Khartoum, Sudan

Salmah Mousbah Zeed Mohammed The School of Computer Sciences, Sirte University, Sirte, Libya

Ikram Mokhammed Department of Therapy and Endocrinology, RUDN University, Moscow, Russia

G. A. Montoya Universidad de los Andes, Bogotá, Colombia

Yoshitatsu Mori Osaka Electro Communication University, Osaka, Japan

Mikram Mounia Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

Nompumelelo Mthethwa Audit and Taxation, Audit and Account Management Department, Durban University of Technology, Durban, South Africa

Seongmi Mun Seanus Co, Busan, Republic of Korea

Miranda Mundt Sandia National Laboratories, Albuquerque, NM, USA

Takayuki Nakatani Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Purab Nandi Department of Electrical and Electronics Engineering, BITS Pilani, Zuarinagar, Goa, India

Mitchell Negus Sandia National Laboratories, Albuquerque, NM, USA

Bethany L. Nicholson Sandia National Laboratories, Albuquerque, NM, USA

Hiroshi Noborio Osaka Electro Communication University, Osaka, Japan

Doruntina Nuredini University of Information Science and Technology “St. Paul the Apostle”, Ohrid, Macedonia

Mızan Islami Nurzihad Master of Law, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Andreas Oberweis FZI Forschungszentrum Informatik, Karlsruhe, Germany

Toshiyuki Okamoto ROHM Semiconductor, Yokohama, Japan

Atsushi Okuyama Tokai University, Hiratsuka, Kanagawa, Japan

Inara Opmane Institute of Mathematics and Computer Science, University of Latvia, Riga, Latvia

Jorge O. Ordoñez-Ordoñez Universidad Politécnica Salesiana, Cuenca, Ecuador

Nicholas Ouellette University of New Brunswick, Fredericton, Canada

Nabila Ounasser IMS Team, ADMIR Laboratory, Rabat IT Center, ENSIAS, Mohammed V University, Rabat, Morocco

Se-Hoon Park Korea Orthopedics and Rehabilitation Engineering Center, Incheon, Republic of Korea

Pilca Imba Wilmer Patricio Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Nikolay Pavlov Technical University of Sofia, Sofia, Bulgaria

John Pender United States Department of Agriculture, Washington, DC, USA

Markus Peterhansl Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung, Germany

Brayan F. Peñafiel-Pinos Universidad Politécnica Salesiana, Cuenca, Ecuador

Rainer Poeschl Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung, Germany

Nanik Prasetyoningsih Master of Law, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Ulung Pribadi Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Budhi Prihartono Faculty of Industrial Technology Institut Teknologi Bandung (ITB), Bandung, Indonesia

Michele-Luca Puzzo University of Rome (Sapienza), Roma, RM, Italie

Zuly Qodir Department of Islamic Politic – Political Science, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Praveen C. V. Raghava Bosch Global Software Technologies, Bangalore, India

Moli Aya Mina Rahma Master of Law, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Adnan Rashid School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan

Lilia Raycheva The St. Kliment Ohridski Sofia University, Sofia, Bulgaria

Md Tanzim Reza BRAC University, Dhaka, Bangladesh

Maryem Rhanoui IMS Team, ADMIR Laboratory, Rabat IT Center, ENSIAS, Mohammed V University, Rabat, Morocco;
Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

M. Ragib Anjum Riad BRAC University, Dhaka, Bangladesh

Mazza Rocco University of Campania "Luigi Vanvitelli", Caserta, Italy

Siegfried Roedel Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung, Germany

Jonathan Francis Roscoe BT Applied Research, Adastral Park, UK

Miguel Zenon Nicanor L. Saavedra Ateneo de Manila University, Quezon City, Philippines

Bidyut Saha Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung, Germany

R. K. Saket Department of Electrical Engineering, Indian Institute of Technology (BHU), Varanasi, UP, India

Okfalisa Saktioto Informatics Engineering Faculty, Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim, Riau, Indonesia

Andrea Sánchez-Zumba Pontificia Universidad Católica del Ecuador Sede Ambato (PUCESA), Ambato, Ecuador;
Universidad Técnica de Ambato (UTA), Ambato, Ecuador

Renato Santos Department of Informatics Engineering, University of Coimbra, Centre for Informatics and Systems of the University of Coimbra, Coimbra, Portugal

S. S. Sarangdevot Department of Computer Science and Information Technology, JRN Rajasthan Vidyapeeth (Deemed to Be) University, Udaipur, Rajasthan, India

Hasrini Sari Faculty of Industrial Technology Institut Teknologi Bandung (ITB), Bandung, Indonesia

Dipti K. Sarmah Services and Cyber Security Department, University of Twente, Enschede, The Netherlands

Keno Sato ROHM Semiconductor, Yokohama, Japan

Yuki Sato Department of Mathematical Science and Electrical-Electronic-Computer Engineering, Akita University, Akita, Japan

Azali Saudi Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia

Leizheng Shu Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Haidian District, Beijing, China

Shamila Mohamed Shuhidan School of Information Science, College of Computing, Informatics and Media, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

Sergey Shulga V. N. Karazin Kharkiv National University, Kharkiv, Ukraine

Riaz Uddin Shuvo Code Optimizer, Dhaka, Bangladesh

Muhammad Ehsan Siddique ISEP—Institut Supérieur d'Électronique de Paris, Paris, France;
Université Paris Cité, Paris, France

Yousfi Siham Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

Paula Alexandra Silva Department of Informatics Engineering, University of Coimbra, Centre for Informatics and Systems of the University of Coimbra, Coimbra, Portugal

Kuldeep Singh National Institute of Technology, Kurukshetra, Haryana, India

Vikram Singh National Institute of Technology, Kurukshetra, Haryana, India

Max Smith-Creasey BT Applied Research, Adastral Park, UK

Insu Song College of Science and Engineering, James Cook University, Singapore, Singapore

Halima Soussi Laboratory of Applied Geophysics, Geotechnics, Engineering and Environmental Geology (L3GIE), The Mohammadia School of Engineers, Mohammed V University, Rabat, Morocco

Daniel Staegemann Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Martin Stoev Services and Cyber Security Department, University of Twente, Enschede, The Netherlands

Andreas Stöckl University of Applied Sciences Upper Austria, Hagenberg, Austria

Aditya Subash College of Science and Engineering, James Cook University, Singapore, Singapore

Sujith Nyarakkad Sudhakaran Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Girish Sundaram University of Arkansas at Little Rock, Little Rock, AR, USA

Misaki Takagi Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Praowpan Tansitpong NIDA Business School, National Institute of Development Administration (NIDA), Bangkok, Thailand

Kexin Tao College of Science and Engineering, James Cook University, Singapore, Singapore

Chan Eang Teng Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur, Malaysia

Savio Jojo Thalakkotoor Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Anna W. Topol IBM Research—Watson, New York, USA

Philane Tshabalala Central University of Technology, Bloemfontein, Free State, South Africa

Klaus Turowski Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Aigerim Ualihanova Department of Therapy and Endocrinology, RUDN University, Moscow, Russia

Abraham Usman Usman Department of Telecommunications Engineering, Federal University of Technology, Minna, Niger State, Nigeria

Vassilios G. Vassilakis University of York, York, UK

Neli Velinova The St. Kliment Ohridski Sofia University, Sofia, Bulgaria

Aanchal Verma Department of Electrical Engineering, Indian Institute of Technology (BHU), Varanasi, UP, India

Pauline Andrea Vivero School of Information Technology, Mapúa University, Manila, Philippines

Lukas Walker ETH Zurich, Zurich, Switzerland

Xue Wan Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Haidian District, Beijing, China

Wenjun Wang Wenzhou-Kean University, Wenzhou, China

Alexander Weinberger Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung, Germany

Maria Weinreuter FZI Forschungszentrum Informatik, Karlsruhe, Germany

Youngeun Wi Sookmyung Women's University, Seoul, Republic of Korea

Akihiro Yamamura Department of Mathematical Science and Electrical-Electronic-Computer Engineering, Akita University, Akita, Japan

Albashir A. Youssef Arab Academy for Science, Technology and Maritime Transport, Cairo, Egypt

William Emmanuel S. Yu Ateneo de Manila University, Quezon City, Philippines

Artem Yurovsky Clinical Hospital of Civil Aviation, Moscow, Russia

Bissera Zankova "Media 21" Foundation, Sofia, Bulgaria

Mourad Zegrari Structural Engineering, Intelligent Systems and Electrical Energy Laboratory, Ecole Nationale Supérieure des Arts et Métiers, University HASSAN II, Casablanca, Morocco

Bihui Zhang Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Haidian District, Beijing, China

Yujie Zhao Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, Maebashi, Japan

Hafida Zrouri National School of Applied Sciences, Mohammed Premier University, Oujda, Morocco

Santiago Zurita-Armijos Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Overlay Robotized Datacenter System

Khaled Elbehiery and Hussam Elbehiery

Abstract For decades, the world's most valuable known natural resources are oil, radioactive materials, and gold, which all lead to wealth. However, technology has become more pervasive in every aspect of our lives and a new unnatural resource has emerged and becomes worth more; data. Over the past few years, the exponential growth of data is not correlated with the datacenters to handle and process these amounts of data. Building new datacenters takes years to complete because they heavily depend on humans to build, deploy, maintain, protect, and operate all related assets. Most advanced datacenters' designs are embedding robotic arms and technological tools that might be the solution for the future. These designs have their own caveats, and most importantly, they will not fix the problems that we are experiencing today. What the world needs is not to wait for the future and to take a step forward toward the robotized datacenters today to be able to close on the gab of the technical demands and also to support the future designs without the need to rip and replace.

Keywords Datacenters · Robotized datacenters · Hyperscale datacenters · Greenfield deployment · Brownfield deployment · Overlay robotized datacenter system

1 Introduction

The proposed design "Overlay Robotized Datacenter System" introduces a complementary addition solution to the existed datacenters that already built (Brownfield deployment) and the new install datacenters (Greenfield deployment) to accommodate to the consumer demands, and it is also a step forward to support the next generation of datacenters in the future. In turn, it reduces the capital expenditures

K. Elbehiery (✉)
Devry University, Denver, CO 80126, USA
e-mail: khaled.elbehiery@gmail.com

H. Elbehiery
Vanridge University, Denver, CO 80126, USA

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_1

(CapEx) and operating expenses (OpEx) and increases the operation's efficiency and accuracy.

The research paper begins with an overview of the datacenter's technology, which covers the efforts to build datacenters, the surrounding environmental and energy growing demands, policies, and regulations. The research paper discusses how robotic technology is helping many industry fields to grow, and it is accelerating the global and financial economy.

The research paper explains the current problems and issues of building and expanding the datacenters and also the concerns with the futuristic trials of datacenters. The paper expands on the proposed solution "Overlay Robotized Datacenter System" from an architectural view to operating it and the business values of the introduced design. Finally, the research paper ends with market analysis of the proposed design that makes it appealing to enterprises to adopt.

2 Datacenter Technology

A few decades ago, the ordinary way to run an application and store any data was through a personal computer. But over time with the exponential growth of technology in many fields, a centralized dedicated building or group of buildings to house bigger computing systems to serve millions if not billions of users was the solution; this is called a datacenter. Datacenters require certain demands to operate properly such as power, cooling system, racks, airflow, fire protection system, security, and most importantly trained knowledgeable personnel.

Datacenters' capacity varies depending on the services they offer to the consumers. Public cloud providers' datacenters (Amazon AWS, Microsoft Azure, Google GCP, and others) are by far the biggest in the world, and they are considered hyperscale datacenters (see Fig. 1). Hyperscale is increasingly used to define not just the scale and size of these datacenters, but also their architecture. Hyperscale data centers have a minimum of 5000 servers and at least 10,000 square feet in size. Beyond the footprint and server figures, equally important is what is going on inside, where they are architected for a homogeneous scale-out Greenfield application portfolio using increasingly disaggregated, high-density, and power-optimized infrastructures.

IoT, 5G, and artificial intelligence (AI) technologies along with faster chipsets are increasing the demand for computing capacity and in turn infrastructure and energy. Unfortunately, the time to build those facilities and get them ready to accommodate to the technology and business demands are not in correlation, and this gap is constantly growing even considering all energy resources in the future. Also, another reason for slowing down the progress is environmental regulation such as net zero initiatives and carbon footprint reduction policies (see Fig. 2) [1].

Fig. 1 Hyperscale datacenters

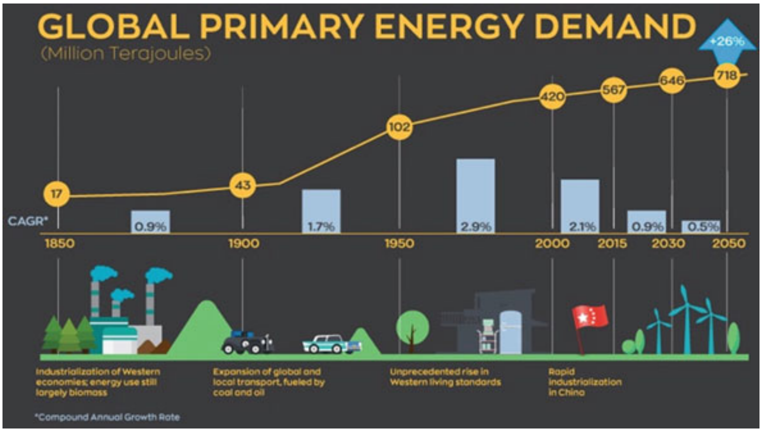


Fig. 2 Global energy demands

3 Robotic Industry

Manufacturing methods have been heavily dependent upon manual labor and skills; today, the programming work, the monitoring, and the calibration are all automated and done through the computer such as repetitive jobs in factories, sewing machines, painting, and more. It is a collaborative effort between humans that have the intelligence, instincts, and reflexes and the machines that have the advantage of doing things faster and with a lot of precision.

The development of intelligent robots helps humans to achieve things that currently seem impossible such as dangerous or daring environments. Artificial intelligence becomes inseparable from robotic engineering. They are together a disruptive technological approach and the next breakthrough in numerous technical fields such as automated factories, economy, and transportation but also in non-technical fields, such as health care, disability, and finding cures for complex brain diseases [2].

The most important thing to remember is that any technology should be invented in a way to help humans with their capabilities, disabilities, and needs. An application of robotic engineering based on artificial intelligence (AI) is the exoskeleton; the exoskeleton is a candidate to help out where humans are having to do physically demanding work such as construction, manufacturing of automobiles and planes, and warehousing.

Technology is advancing rapidly to help with solutions to problems and improvements to our lives. According to the International Federation of Robotics (IFR) study World Robotics 2018, there were about 2,097,500 operational industrial robots by the end of 2017. This number is estimated to reach 3,788,000 by the end of 2023. For the year 2017, the IFR estimates the worldwide sales of industrial robots with $16.2 billion. Including the cost of software, peripherals, and systems engineering, the annual turnover for robot systems is estimated to be $48.0 billion in 2017 and is growing ever since [3].

China is the largest industrial robot market, Japan had the largest operational stock of industrial robots, and in the USA, the industrial robot makers' shipping rates to factories are accelerating exponentially [4].

The biggest customer of industrial robots is automotive industry with 33% market share, then electrical/electronics industry with 32%, metal and machinery industry with 12%, rubber and plastics industry with 5%, and food industry with 3% in textiles, apparel and leather industry [5].

In summary, humans have unlimited ambitions for the future, and the fact is the future has not been decided yet, and definitely no limit to how far humans' dreams can reach. Companies such as SpaceX with an ambitious plan to send an unmanned capsule to Mars. The Boring company is digging a vast network of underground tunnels that will change transportation forever. Neuralink Corporation is tapping into the human brain to cure diseases, all that were dreams not too long ago but today they become facts [6].

4 Technology Concerns

Building and operating a datacenter is a significant task that must adhere to many rules and policies, some are environmental and some are technological standards, and the latter focuses on deploying different equipment and associated gear in the standard 19-inch equipment rack such as power units, power cables, copper and fiber optic cables, servers, networking devices, and more. Although the rack dimensions (Length X Width X Height) are standard and there could be many standards, the equipment themselves is changing rapidly over the past few years due to the competitive environment among hi-tech companies which causes the old products to be outdated and new ones to take place. In addition, unfortunate event such as the recent pandemic situation has its own effect globally, especially when it comes to chain supply and demand, and in turn, services are falling behind.

Building bigger size datacenters and migrating services to them should achieve maximum productivity to accommodate to the accelerated demands, but this solution certainly is significantly expensive, consumes years to complete, and it continues to have drawbacks or caveats, a road full of bumps and obstacles, the following are just a glimpse of the major ones:

a. The appropriate locations and the environmental surroundings such as cold weather, rivers, or water supply are not easy to find and financially is very expensive.
b. Millions of applications for billions of customers will be impacted due to the migration to new datacenters, not to mention it is financially a huge burden.
c. The optical fiber network that already connects the current datacenters' locations comes with a very significant cost to replace if possible.
d. Global event such pandemic situation (COVID-19) has already affected the world, and it was not an easy lesson and confidently to say, the world is not ready yet for another one.
e. Governmental services that overrule technological demands.
f. The ever-increasing compliance demands placed on facility managers today.
g. Recruiting, hiring, and training new crews at the new locations do not happen overnight, and usually the majority of workers after many years do not likely to move.

The next generation's datacenters are supposed to be fully automated, but this is a very difficult thing to achieve, and the fast pace of technology movement has vastly impacted the business worldwide. Unfortunately, the challenge is that many sites and places are not completely ready yet to adapt to this technological movement. The following also represent some of the caveats that are found with the fully automated hopeful designs of datacenters:

a. The futuristic designs proposed are likely to discuss a very special case that is dealing only with a new build datacenter (Greenfield deployment) with specific infrastructure and floor plan, a specific kind of racks, and the connectivity is done in a very specific way for specific kind of equipment.
b. Adapting and programming the robotic system to different kinds of datacenters' equipment due to technology's accelerated development is not an easy task, and it will consume time, effort, and money if this is even possible.
c. The futuristic automated datacenters' designs fit special cases of business and technology models such as public cloud providers (Amazon Web Services (AWS), Google Cloud (GCP), and Microsoft (Azure)) which is based upon having unlimited computational resources (compute, memory, and storage). These computational resources definitely could be stacked by type in dedicated aisles in the datacenter facility to serve millions of customers which in turn makes a specific kind of automation a feasible option. Very important to remember that there are numerous companies that have their datacenters in totally different

deployment fashion which is by far considered the dominant, unlike the public cloud providers' as a special case.

5 Overlay Robotized Datacenter System

5.1 Architectural Design

The "Overlay Robotized Datacenter System" design is fundamentally based on deploying a commercial robotic system in the current datacenters. The robotic mechanism would be able to move around between the rack-mounted equipment in the aisles of the datacenter's floor. The robotic arms would come with different types of replaceable accessories for different purposes. Figure 3 shows an overview of deploying the proposed robotized system in the datacenter' halls and aisles.

The robotic arm system could be a single unit per aisle's rail to handle a task for equipment deployment or a dual unit fashion to handle multiple functions simultaneously on one aisle's racks (see Fig. 4).

The robotic arm can work in an "Independent Mode" to serve different tasks or functions for any rack in the aisle. It also can work in a "Join Mode" which works with the next aisle's robotic arm coherently to serve a task that requires a joint effort of the two robotic arms, such as installing a heavy weight equipment. This kind of work ordinarily required two to four datacenter's individuals to accomplish (see Fig. 5).

There are three primary categories and are not limited to what kind of accessories the robotic arm could be equipped with, the robotic arms accessory depends on the

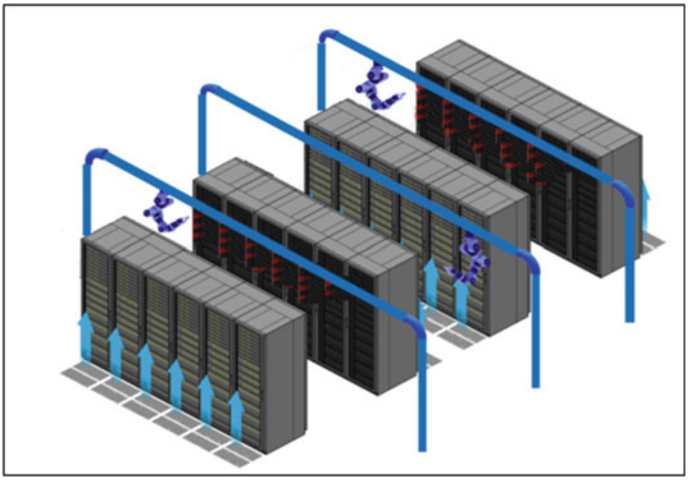


Fig. 3 Overview of the proposed robotized datacenter system

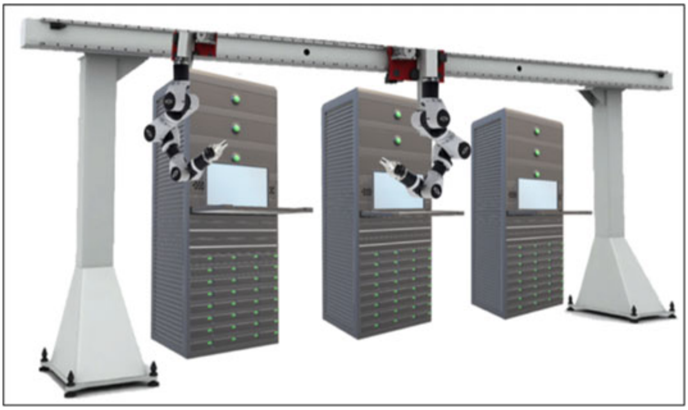


Fig. 4 Robotized arm system deployment

Fig. 5 Robotic arm operation modes

required purpose or the function that is needed. Bear in mind that all accessories are replaceable to one another to serve any goal anytime; the following are some examples:

a. Heavy weight lifting accessories that help with installing racks, power, and cooling units.
b. Lightweight lifting accessories that help with installing servers and network devices.

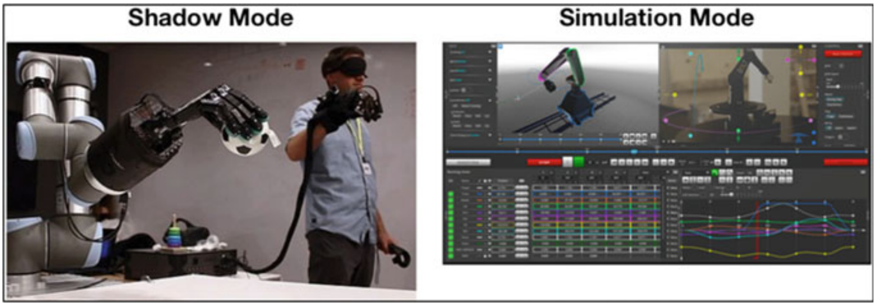


Fig. 6 Robotized system motion modes

c. Operations and support accessories such as environmental sensors, camera, security/surveillance, and remote work/telexistence)

Overall, the "Overlay Robotized Datacenter System" motion has different modes of operations:

a. ***Manual Mode***: The robotic arm's motion, movement, and function could be controlled wirelessly from a GUI interface integrated application on a handheld device such as an iPad or a tablet.
b. ***Auto Mode***: Some tasks such as observation, surveillance, and more do not need human intervention and could be set up in an automatic fashion.
c. ***Shadow Mode***: The robotic arm's motion follows the datacenter's individual motion that in turn comes as a handy option for remote control from the datacenter's control center.
d. ***Simulation Mode***: This mode is designed for testing, calibration, and fine-tuning functions for tasks before moving them to an operational phase for safety purposes (see Fig. 6).

The datacenter operation control center has the complete vision of the robotized datacenter, and the control center has full access to all the telemetries' data from the robots' arms such as cameras and sensors and the accurate positions as well (see Fig. 7).

Any human presence in the datacenter along with the exact location/coordinates is detected immediately to the control center which in turn overrules and disables any automatic tasks for all the robotic arms for safety purposes, in addition to the standard tasks of the control centers such as security and surveillance that create opportunities for savings across all functional areas (see Fig. 8).

5.2 *Business Values*

In order for a company to maintain its leadership in the innovation of new attributes, it must learn to offer product innovations routinely which is going to lead to lower

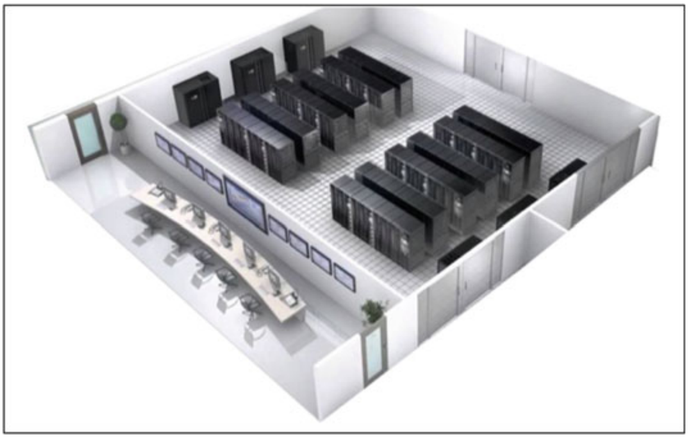


Fig. 7 Overview of robotized datacenter operation control center

Fig. 8 Robotized datacenter operation control center safety and surveillance

the prices and foster the development of new technology. The significant advantage of using the "Overlay Robotized Datacenter System" is that despite it might have an upfront cost, it is faster to serve the broad range of datacenters' facilities today and in the future. Only for unique cases when the business owner is occupying a floor in a facility and operates it as datacenter instead of a dedicated facility, then adding the new "Overlay Robotized Datacenter System" will require extra cautions to implement considering the other business in the same facility safety.

The "Overlay Robotized Datacenter System" is totally a new innovative architecture product, and no attempts have been made to manufacture it or market it. It is designed for datacenters' facilities that are already in service (Brownfield deployment) as well as new build datacenters' facilities (Greenfield deployment). It is a more agonistic approach that works for many purposes and for different tasks. It works today, tomorrow, and in the next decade to satisfy a long-term Return of Investment (ROI).

The robotic system has the option to support various accessory devices or could also be equipped with different kinds of environmental monitoring such as sensors and cameras for the purpose of maintaining visibility into the datacenter's environment or remote work which in turn will have a huge reduction in downtime, troubleshooting, and surveillance.

The robotic arms should be able to handle different kinds of weights, from heavy lifting such as assembling and disassembling data center racks, power, or cooling system to servers or any kind of equipment to be deployed inside the racks no matter what their configuration's standards are.

The "Overlay Robotized Datacenter System" is depending on the human supervision presence that has core skill set, and it will always be there for decades to come for the following reasons:

a. To fill in the gaps and supervise in non-fully automated fashion system or tasks.
b. To empower the artificial intelligence (AI) system with more details about the required tasks to be able to automate some of those tasks eventually.
c. To control manually or remotely the robotized system or work jointly with it.

It keeps unemployment rate very low, and it provides a leap jump to the future through accelerating the technological level of the datacenter's employees. It offers opportunities to another class of employees; the handicap individuals who do not exist today in this specific work environment.

The "Overlay Robotized Datacenter System" solution can learn how to become more adaptive, more efficient, and substantially automate labor-intensive tasks which contributes to avoiding or reducing the human injury, stress, and fatigue to the minimum.

It fits any datacenter's building infrastructure, it works with any kind of racks, and it can deploy any type of equipment that are ready to be installed inside the racks. It is a multi-purpose system and executes different tasks, not tied to any sort of a company or proprietary device or product.

The "Overlay Robotized Datacenter System" process can be simulated before actual operation to save time and increase the level of safety associated with robotic equipment. The ability to preview the behavior of a robotic system in a virtual world allows for a variety of mechanisms, devices, configurations, and controllers to be tried and tested before being applied to a "real-world" system.

It is worthy to note that the production of an automated manufacturing system would likely need to meet a wide range of quality production requirements such as the International Standards Organization (ISO) as well as electrical and mechanical safety standards under the Occupational Health and Safety Administration (OSHA) products using digital electronics would also be subject to various regulatory requirements including those of the Federal Communications Commission (FCC). Other regulatory bodies which could influence the manufacture of the proposed "Overlay Robotized Datacenter System" may include, but not necessarily be limited to, the International Electrotechnical Commission (IEC) and the International Telecommunications Union (ITU).

5.3 Market Analysis

Five major technological growth factors are foreseen to rabidly drive the demands of the "Overlay Robotized Datacenter System":

a. Building automation systems will continue to grow in usage, as they are centralized, interlinked networks of hardware and software that monitor and control the environment in commercial, industrial, and institutional facilities.
b. Governments of various countries are adopting regulations to minimize energy usage and waste.
c. New technologies driving global market demand include web-based or cloud-based control systems supported by IoT, wireless and mobile technologies, integrated building system, and facility management solutions.
d. Security system integrated with other building systems creates opportunities for savings across all functional areas.
e. Advanced data analytics on cloud-based platforms has opened up a whole new world for cost savings and operational efficiencies, giving facility managers the ability to make their buildings smarter and more intelligent over time.

Security and access controls currently account for the majority of revenue share in the global building automation systems' market. It is anticipated to hold the largest market share of more than 30% throughout the forecast period, including solutions for safety-critical services (e.g., fire or security alarm systems) and security-critical services (e.g., intrusion alarm or access control systems).

Today, companies are utilizing automated ground vehicles, robotics, and automation within their manufacturing and industrial facilities to realize great savings. Companies are also starting to use drones to patrol outdoor property, saving money on guards and manpower, and minimizing risks and errors by humans.

The primary market for the proposed design should target the owners of the datacenters. As of toward the end of year 2022, there are more than 750 hyperscale data centers in the world. By 2026, it is estimated they will get to 1200 (see Fig. 9) [7].

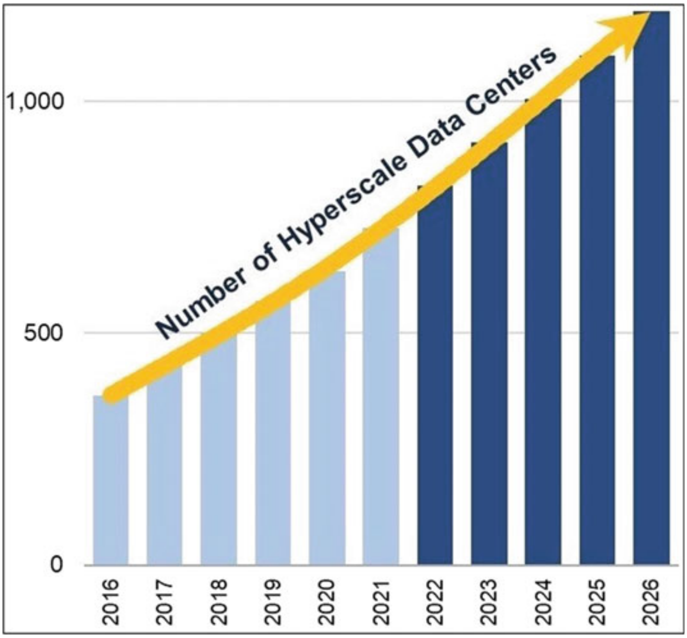


Fig. 9 Datacenters worldwide

6 Conclusion

Technologies that are surrounding our lives such as IoT, 5G, artificial intelligence (AI), and machine learning/deep learning (ML/DL), along with companies like Amazon, Microsoft, Google, Facebook, and Netflix, are the engines behind the new data economy, and it continues to expand globally and exponentially.

Simultaneously, robotic technology is literally invading almost every industry today, robotic arms have replaced humans on the factory floor, and they perform pre-programmed repetitive tasks much more reliably than humans. The robotic arms are not only providing accurate results more than any human being is capable of but also the capability of doing tough jobs such as lifting weights with no fatigue, no lunch hours, not going home, and not only working eight hours per day like humans. Very important to remember that despite the labor jobs are being replaced by robots, other jobs are taking place such as programmers, observers, quality controls, assurance, and supervisors. Technology did not look back, but it combined artificial intelligence (AI) with robotic automated systems that have advanced rapidly to almost every aspect of our lives, [8].

All in all, breakthroughs continue to happen each and every day in all the technology fields around the world. The amount of data surrounding these technologies has exceeded expectations and consequently that presents multiple challenges for the organizations and infrastructure required to support it. Building new datacenters using the traditional methods of today might seem to be the solution, but it has a price tag such as extreme high cost, significant effort, and years to complete. On the contrary, companies are working on experiments and trials for the future and they are definitely reasonable solutions, but that will need from us to wait years from today until the new design is available to implement and perhaps solve the problems. The realistic approach is to have a design that it can work properly today and support future needs as well.

The "Overlay Robotized Datacenter System" is designed for datacenters already in service (Brownfield deployment) and the new install datacenters (Greenfield deployment) as well through deploying commercial robotic system that could come with different types of mobile replaceable arms along with the moving mechanism between datacenters' aisles [9].

The proposed solution reduces deployment expenses and time consumed while simultaneously increases efficiency through a cohesive integration between the intelligent human operator and the commercial robotic system that is capable and adaptive to future demands. The whole system could be fully operated efficiently by people with disability, which in turn opens up a great opportunity in a technology field that is dominated by ordinary healthy individuals, and it improves task force management by lowering the unemployment rate as well [10].

The "Overlay Robotized Datacenter System" reduces the time to deploy equipment, which expedites the time to offer services to consumers. It significantly increases employees' safety by introducing a safety hazard environment against contamination that could endanger humans. Most importantly, it is highly compatible to work with severe conditions such as pandemic, epidemic, or outbreak health situations.

References

1. Born to Engineer Blog (2018) Is engineering the solution to global energy demand?. Infographics, USA. Available online: https://www.borntoengineer.com/is-engineering-the-solution-to-global-energy-demand
2. Elbehiery K, Elbehiery H (2020) Millennium robotics; powered by artificial intelligence and cloud engineering. Int Organ Sci Res (IOSR) 10(04):44–53, Series II, ISSN (e): 2250-3021 ISSN (p): 2278-8719, (2020). Available online: http://iosrjen.org/pages/volume10-issue4(series-2).html
3. International Trade Administration, U.S. Department of Commerce (2015) U.S. Export fact sheet. USA. Available online: https://www.trade.gov/
4. Daniel Workman (2016) United States Top 10 Exports, USA. Available online: https://www.worldstopexports.com/united-states-top-10-exports/
5. Wood L (2017) Global $100+ billion building automation market forecasts 2017–2022—research and markets. Cision Distributionm, USA. Available online: https://www.prnewswire.com/news-releases/global-100-billion-building-automation-market-forecasts-2017-2022---research-and-markets-300432738.html
6. Elbehiery K, Elbehiery H (2020) Coronavirus; aftermath technology outburst Seventh Sense Research Group®. Int J Electr Commun Eng (IJECE) 7(6):1–12, Louisiana, USA, P-ISSN: 2349-9184, E-ISSN: 2348-8549. https://doi.org/10.14445/23488549/IJECE-V7I6P101

7. Maistre RL (2022) How many hyperscale data centres does the world need? Hundreds more, it seems. TelecomTV, © Decisive Media Limited 2022, UK. Available online: https://www.telecomtv.com/content/digital-platforms-services/how-many-hyperscale-data-centres-does-the-world-need-hundreds-more-it-seems-44015/
8. Mid-Atlantic Controls Corp. (MACC) (2017) Intelligent building technology and automation trends for 2018, USA. Available online: https://info.midatlanticcontrols.com/blog/intelligent-building-technology-and-automation-trends-2018
9. Miller R (2020) Will robots usher in the lights-out data center?, Datacenter Frontier, Copyright Endeavor Business Media© 2022, USA. Available online: https://datacenterfrontier.com/will-robots-usher-in-the-lights-out-data-center/
10. Focke R (2017) How integrating security and building automation saves time, energy. Campus Safety (CS) magazine, © 2022 Emerald X, LLC., USA. Available online: https://www.campussafetymagazine.com/contact_us/

Development and Applications of Data Mining in Healthcare Procedures and Prescribing Patterns in Government Subsidized Welfare Programs

Praowpan Tansitpong

Abstract Electronic medical records are crucial for the development of government subsidized programs in modern healthcare management. In this study, data mining techniques are used to identify prescribing patterns in health insurance plans and to determine whether differences between health insurance plans and benefits affect healthcare delivery. Electronic medical records were collected from rural hospitals in Thailand according to National Health Service guidelines. This study shows the cost structure of the Thai government's healthcare program. Due to the variety of drugs and complexity of medical service, the reimbursement cost for patients is much higher in social security programs.

Keywords Healthcare process variations · Service differentiation · Electronic medical records · Health benefit programs · Healthcare data mining

1 Introduction

Launched in 2001, Thailand's Universal Insurance Scheme (UCS) provides medical benefits to about 95% of the total population, in addition to the Civil Servant Medical Benefit Scheme (CSMBS), which covers about 10,000 people. Thailand also provides a social security system with compulsory health insurance for 10 million private sector workers. Thai citizens who cannot afford social security or other private insurance can receive treatment at designated medical facilities. These initiatives require healthcare providers to send electronic data in a structured format to governments, including private and public hospitals. All government hospitals and some private hospitals offer these plans. However, health insurance in Thailand is very complex as the government has created 60 additional social insurance categories for institutional beneficiaries such as local administrators (PAOs), regional administrators (SAOs),

P. Tansitpong (✉)
NIDA Business School, National Institute of Development Administration (NIDA), 118 Seri Thai Road, Bangkok 10240, Thailand
e-mail: praowpan.tan@nida.ac.th

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_2

retired officials, public school teachers, and the families of the beneficiaries. Hospitals in Thailand provide reimbursement to hospital staff and their families as a form of alternative health care. By combining Thailand's main pension schemes (social security and civil service, private insurance, and universal insurance) with hospital health insurance for hospital workers, the health system management system addresses very specific longitudinal differences in medical procedures. Differentiation of services is common in the healthcare sector. Many hospitals and clinics offer service options that allow patients to stay in luxurious private rooms with premium bedding, have their own on-site chef, or receive a private visit. In some hospitals, older patients (also known as "ward patients," or patients who pay extra for past services) are marked in red in their medical records, while inpatients are often marked in white (New York Times, 2015). Premiums and benefits for each benefit plan vary depending on the course of treatment, the specific disease, and the healthcare facility. This practice is also occurring in most countries globally.

The literature on electronic big data in Thailand's healthcare system is limited due to the complexity of data collection and the multiple structure of databases for different diseases, personal data, and government audits. Health information is extensive, but public and private insurances also depend on the savings account and the patient's plan and options. Generic health insurance does not cover horizontal and vertical difference in treatments and prescriptions. The benefits are not due to the full vertical difference between private and public insurance, because taxes are less selective than private savings and patients have to pay more for insurance in public projects such as Medicare. Since the transformation of Thailand's healthcare system in 2000, there has been an increasing demand for medical care. Local hospitals in Thailand have access to local patients due to various government welfare programs. In the medical field, many hospitals use traditional database management, but many electronic medical records (EMRs) are collected. Therefore, most regional hospitals are working to increase the efficiency of their clinical operations by leveraging big data and choosing the right technology. EMRs encourage patients, physicians, nurses, and others to participate in drug therapy monitoring and management. The biggest challenge for Thai healthcare is to reduce costs and improve treatment processes and outcomes. Thailand considers itself the health center of Asia and is trying to solve this problem by analyzing health data to become a leader in health research in the Economic Community (AEC) and to become center of healthcare system in the Association of Southeast Asian Nations (ASEAN).

From 2010, new government subsidized programs are available in Thailand, patient benefits, including dialysis, are more expensive, and self-funded benefits often include (or exceed) family benefits. Rather than comparing differences in coverage between private and public insurance, this study aimed to examine the unique evidence and differences between benefit plans that may affect substance abuse treatment. To examine the effectiveness of standards of care for different patient segments; hence, healthcare dataset should include a full sample of health insurance companies with a sample (or cross-sectional) and vertical differentiation. In health system, resources are limited, healthcare providers such as physicians spend time interacting with patients and focus on diagnosis with constraints in making decisions. The purpose of this study was to understand how electronic health databases can be used to identify prescription drug decision-making patterns among insurance companies and publicly funded benefit providers.

2 Literature Review

Healthcare providers aim at developing variety of products to fit their market segmentation strategy. Hospitals can use segmentation strategies to increase overall demand and generate more revenue in different segments. Another benefit of market segmentation is that companies can charge higher prices for high-value products or services that customers want [1–5]. Patients pay more for quality care. On the other hand, patients who do not use fee-for-service services do not prioritize care, and providers must allocate resources to prioritize clients according to the terms of the contract they choose. Some patients are more critical and may require more shifts than others, and emergency care is often provided earlier than inpatients. However, selection bias can negatively affect other patients and the overall quality of the emergency department [6–8]. In the review of these literature, physicians are expected to offer alternative treatments, but may not be aware of inherited influences on treatment decisions that are influenced by what they know about the patient's condition. EMRs facilitate health management to determine functional decision-making procedure; however, the usage is a relatively new concept in healthcare [9–14]. Because the course of treatment varies, the quality of care provided by healthcare providers is unknown and its effectiveness has not been proven. There is no evidence of a relationship between quality of care and a specific type of prescribing treatment in product or flow of the service. Further research is then needed on the cost-effectiveness of high-quality healthcare. To test this hypothesis, health data were the primary source to determine whether healthcare delivery affects healthcare quality. Medical records can be used to monitor staff care during treatment to patients. Electronic medical records are structured databases that contain information about patient care. Databases provide access to data that can be used to uncover hidden patterns and relationships in decision making. Although some functional deterministic patterns are difficult to identify or explain in everyday practice, EMRs analysis can provide clinicians with some structural results [15–18]. However, the literature ignores empirical evidence for some differences between brand name and generic drugs for disease-specific drugs.

There are many different brands in the drug group that are prescribed for patients with similar diseases. In this study, the authors examine the fairness of allocation decisions for different patient segments defined by different treatment plans and payers. A study of cross-sectional and cross-sectional health benefits and relationships among various determinants, including recommended doses and other criteria, in addition to previous literature providing age, sex, or initial diagnosis data. This study is the first to review the recommended results of the diagnostic model and point out the specific reasons for the different initial results, including the specific benefits and treatments. This study builds on the literature addressing moral hazard behavior in treatment choice and drug use [19–21]. Physicians will prescribe a more expensive drug or a higher dose if patients have allowance for reimbursement. In some countries, a variety of services, including physician visits, hospital costs, and prescriptions, are governed by government regulations [22–24]. In most cases, physicians in Asian countries prescribe the drug, prescribe, and generate profits for hospitals. Researchers have studied health providers' behavior from different perspectives. Given the differences in drug selection (e.g., high-end brand versus generic), the literature has examined the determinants of physicians' vertically differentiated behavior in prescribing for treatments [25–28]. Government subsidy programs also regulate retail prices, and healthcare providers take favorable actions to distort retail prices and price competition through reimbursement and price controls on wholesale and retail prices [29–33]. While these studies focus on the impact of industry on healthcare practices, this study uses electronic health records to analyze the micro-operational decisions of healthcare providers including a given condition may be prescribed to a patient based on multiple prescriptions. In this study, the authors examine the distribution of patients across different healthcare delivery categories and payment programs.

Studies in the past have pain attention to patients benefit and socially allocation welfare on multiple programs [34–37]. According to the literature, the goal of prescribing is to maximize patient benefit, and physicians make decisions to improve treatment outcomes. Social planners may encourage the use of generic drugs over brand name drugs, but the high cost of approved quality makes generic manufacturers inflexible in price competition. Price and brand differences have made branded drugs a popular option in the system of health care in Europe. There are no studies in the past that take into account the classification of recipes according to the type of ingredients. Previous literature has extensively discussed pharmaceutical companies' decisions to develop horizontal and vertical product differentiation to prevent competitors from accessing products throughout the product life cycle [38–42]. These studies suggest that product differentiation facilitates marketing efforts (e.g., advertising) to persuade physicians to prescribe a particular drug brand. However, other publications often target physicians who prescribe brand name or generic drugs. Changes to drug manufacturers, trade names or brands and discount plans must be explained in the regulations. Because this is the first study to examine the volatility of market segmentation as a factor in healthcare decision making, it reflects different perceptions of quality. Two other control variables may have influenced the drug decision. Medications and medicine costs reimbursed to physicians or hospitals. Hence, the main purpose of this study is to investigate prescribing trends in health plans and to use data mining techniques to determine whether health plans and benefit plans influence drug selection decisions for patients treated similarly in other plans. In addition to current literature suggesting that surgeries are related to age, gender, or prognosis, this study examines the relationship between horizontal and vertical benefits and other choices such as cost, medications, and deductibles. The goal is to identify interactions and explore how many placement decisions are made in diagnostic model and uncover underlying reasons for profit seeking behavior, including specific services and treatments for different benefit plans.

3 Methodology

Data collection methods for this study included data retrieval, transformation, and loading. Disease coding data can play a role in insurance-covered treatments. There are three main systems operated by the government: the security system, the health system for civil servants or public employees, and the universal protection system. The universal insurance plan and two other plans (insurance company and single payer) are included in the data collection process. All personal patient information is decoded and converted to a new identifier such as a patient number ID (or HN). Patients with chronic conditions were matched to the International Index of Diseases and Health Problems for several major chronic conditions (hypertension, cancer, and diabetes). The cleaning process also includes separation of digital blocks and text blocks. Open-source software (MongoDB, Hortonworks, CouchDB, Cloudera, etc.) is required to convert comma separated value (.csv) files to JSON. Amazon Web Services (AWS), a recognized cloud service, was selected to host the data and manage this data ecosystem. The National List of Essential Medicines (NLME) lists essential medicines for the prevention and management of all essential medicines, with a focus on treatment recommendations. Diseases, this list was first submitted in 1972, with the most recent update in 2016. The primary purpose of the NLME is to prevent the unnecessary use of medications and to control the overall cost of prescription drugs. However, the government-controlled program allows three physicians to collaborate on the use of off-label medications when necessary for treatment. Violations can be considered from three perspectives. Positive violations (failure to follow instructions), neutrality or adherence to guidelines, and negative violations (violations of guidelines and other normative actions).

The database consists of 18 separate Structured Query Language (SQL) tables. Based on the results of testing and evaluation during data collection, the program with the best performance in terms of speed of program execution, reliability of services, and availability of the program was selected. Patients with chronic diseases were listed in the publication of the International Classification of Diseases 10th Revision. Missing values and zeros were removed as part of the data cleaning process. Translate the number and units of the specified string. The cleaning process also includes ID number and text element. Open-source software (MongoDB, Hortonworks, CouchDB, Cloudera, etc.) is required to convert comma separated value (.csv) files to JSON, which enables configuration and management of communication between secured Amazon Web Service (AWS) and cloud services. The next step is to use the development environment to write a Java program that compares the number of commas in a table column and compares and calculates the value of a given column with the actual data. This method is designed to export data from

hn	c_phloc	p_r_type	**running_no**
ref_date	ref_time	*vn*	*vn_seq*
s_hn	ph_ref_type	*c_pttype*	no_c_pttype
c_clinic	pat_ref_no	pat_pay	ph_pay_flag
total_item	ph_total_amt	ph_right_amt	have_free_amt
order_by	ph_cancel_by	return_by	ph_cancel_date
ph_cancel_time	p_flag	ph_rec_status	cash_flag
mark_up	social_exp_time	spclinic_flag	send_fexpopd
upd_by	upd_date	upd_time	c_pattype
c_contract	s_doctor	pttype_up1	visit_date
repeat_qty	repeat_days	repeat_med_flag	c_right_type
c_location			

Fig. 1 Example of fields in the database

two programs simultaneously. In this way, the primary key values are calculated and reduced to one column, as shown in Fig. 1. This study also compares other programs because the system is designed as a cloud service platform for accessing and managing Big Data organizations. The process begins with obtaining hospital records and processing patient data. During this process, data is collected in comma separated values (CSV) format. The result of this process is shown in Fig. 2.

4 Results

A descriptive overview of the variables is given in Table 1. The log function was used to generate a uniform distribution of clinically useful parameters when evaluating inpatient and outpatient settings (profit and reimbursements). Outpatient prescriptions have advantages over clinical prescriptions. The mean, standard deviation, and standard deviation of brand preference are similar, indicating that the pharmaceutical company is associated with the drug brand. The maximum number of hospitalization wards (shifts) for this diagnosis. There are indications that overprescribing can be explained by limited changes in the quality of care. These drugs are more common among older people in Southeast Asia. Differences in efficacy, dose, choice of brand, diagnosis treatments, and profit are investigated. The regression model provided predictors for retail value, number of prescriptions, manufacturer diversity, brand preference, and preference. The regression results are shown in Table 1. Differences in dose, brand choice, and benefit plan were significant only between the two models combined. Profit is calculated as the difference between the retail price and the cost

Table Name: diagnosis				Table#1		
Table Type: Master						
Description: -						
Field Name	**Type**	**Length**	**Description**	**Key**	**Reference**	**Null**
c_diag	nvachar	5		PK		
diag_seq	int	4		PK		
c_dept	nvarchar	2				
c_sub_dept	nvarchar	2				
diag_name	nvarchar	250				
diag_name_thai	nvarchar	250				
group_icd	int	4				
valid_sex	nvarchar	1				
valid_fr_age	int	4				
valid_to_age	int	4				
c_icd_grp1	nvarchar	3				
c_icd_grp2	nvarchar	3				
c_icd_grp3	nvarchar	3				
status	nvarchar	1				
upd_date	int	4				
upd_time	int	4				
upd_by	int	4				
comorbidity_flag	nvarchar	1				✓
external_flag	nvarchar	1				✓
select_chk	nvarchar	1				✓
c_diag_como	nvarchar	5				✓

Fig. 2 Diagnostic query results

of the drug. Inpatient services (IPDs) were separated from outpatient prescriptions (OPDs), requiring decision making for different patient groups.

A regression model describes relationship between benefit plan and variation in treatment procedure, amount of prescription, drug choice, dose, and manufacturer, all equal. The dose prescribed will depend on the medical condition, but the physician may increase the dose for maximum benefit. Additionally, the results suggest that hospitals may benefit more from a variety of brands and payment methods, but the results are inconclusive. Physicians try to prescribe more expensive treatments whenever possible. For government suggested regulations or other restricted options, on the other hand, limit the drug options and limit the ability to use healthcare after the

Table 1 Descriptive summary

Variable	Obs	Mean	Std.dev.	Min.	Max.
Profit	34,262	31.69169	223.9225	0.3	6107
Reimbursed cost	34,262	257.076	2470.474	0	26,393
Amount of prescription	34,262	61.07184	113.924	0.1	600
Number of brands	34,262	164.8355	78.95719	1	327
Number of manufacturers	34,262	149.8599	67.32339	1	290
Number of multiple cases	34,262	12.2057	5.483112	1	18

option expires. Brand choice has a huge impact on a hospital's profitability. The more brands physicians choose and the more competitive they are, the more profitable the hospital is. The reward index and the acceptable size factor are estimated positively. As a result, the higher the government subsidy, the higher the hospital's prescription income. Hospitals can also make more profit because physicians prescribe higher doses for patients. The higher the government subsidy, the higher the hospital's profit per prescription. Hospitals can also benefit greatly when physicians prescribe higher doses. The results suggest that differences in payment patterns have a positive effect on payment and that hospitals rely on prescribing patterns when physicians have multiple types of patients with specific diseases (Table 2).

Table 2 Regression results of impact of cost structure on healthcare providers' profit

	Profit
Reimbursed cost	0.887***
	(0.00179)
Number of manufacturers	0.041***
	(0.0000585)
Number of multiple cases	− 0.072***
	(0.000162)
Amount of prescription (DOSAGE)	0.466***
	(0.0000366)
Number of brands	0.098***
	(0.0000557)
N	34,262
R2	0.749
adj. R2	0.749
F	4636.1

Standardized beta coefficients; Standard errors in parentheses * p < 0.05, ** p < 0.01, *** p < 0.001

5 Discussion

Since the physician has several options to choose from prescribed doses are measured in pharmaceutical units, the dose may vary according to the physician's prescription. The dose is also adjusted according to similar diagnostic procedures and other brands of drugs. The number of brands is based on the number of different drug names available in the International Classification of Diseases (ICD-10) which includes disease codes, signs and symptoms, pathological findings, complaints, social conditions, and external causes of injury or disease. Brand selection refers to a variety of products, including brands and generics from many manufacturers. Hospitals can work strategically to improve different outcomes for both inpatients and outpatients based on length of stay and treatment required. For example, cost–benefit analyses between hospitalizations in these two groups were evaluated separately. Prescription drug costs are covered by reimbursement and payment plan according to the employed benefits. Although the government sets different reimbursement rates for each prescription drug and hospitalization, the prescription reimbursement system is not always linked to the benefit system.

The number of manufacturers is calculated based on the number of pharmaceutical companies in an ICD-10 diagnostic table. Pharmaceutical companies can play an important role in marketing and advertising expenditures that can affect drug programs. These parameters are related to market openness of pharmaceutical companies and competing treatments for certain diseases. Health insurance changes are the number of health insurance policies that change due to a single diagnosis. In many cases, other patients with similar pathologies are also involved. Age-related diseases such as diabetes are common among retired civil servants and their families. Changes to the unemployment benefits system mean that many patients are moving away from the solutions prescribed by physicians. Dosage gives the physician (or hospital) a different picture of the drug's quality and cost, which can influence prescribing choices. Although this study did not compare differences between private and public endowments, it did look at unique benefit plan adjustments and changes that could affect administrative processes. Reimbursement costs for each benefit plan vary by type of treatment and disease.

6 Conclusion

This study investigated prescribing variability and chronic disease symptoms in diagnosed patients by mapping prescribing patterns in electronic health records to better understand the effects of prescribing changes on patients. The standard treatment procedures and prescription patterns of the Thai government's Supplementary Support System (UCS), Social Security System (SSS), Civil Servant Medical Benefit Scheme (CSMBS), and private insurance are quite different. The analysis also identified inconsistent prescribing patterns in the clinical database. This study

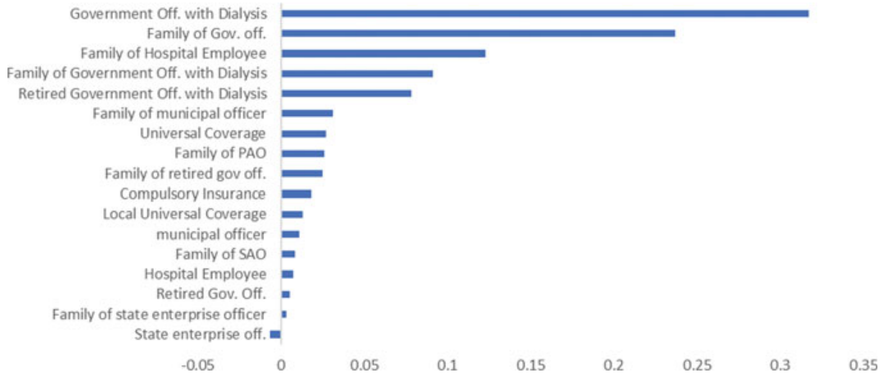


Fig. 3 Government subsidized program profit versus loss comparison

examines differences between benefit plans that may affect treatment outcomes when multiple options are available, which in turn may affect the profitability of healthcare providers. This study shows how big data and cloud technologies can be integrated in practice and how the transfer of data from local hospital databases to the cloud can be supported. This study focuses on database technologies that can improve the performance of medical software and ensure the delivery of medical services. In the eastern part of Thailand, diagnosis data (ICD10) and prescriptions and prescription guidelines from health authorities were collected. As a result, it was found that the co-payment did not affect the cost of treatment, but the characteristics of the frequency of prescription reflected the difference in the service system. Most prescriptions were selected, but some prescription products did not meet the criteria as shown in Fig. 3. Results of this study represent the difference between the compensation plans and the greater the difference between the compensation plans, the higher the rate of return.

In particular, the results show that the differences in benefits and treatment decisions in terms of doses are co-existed. The results are consistent with existing observations that hospitals can build benefit systems based on price structures. The EMRs are essential for physicians and healthcare facilities to make decisions and treat patients. Large amounts of unmanaged data are not available simultaneously. Applying this knowledge of database management will be helpful in solving healthcare process problems and leveraging scarce resources in healthcare management. This helps both physicians and staff to work more efficiently and provide the best care or service to patients. The results also show that brand selection and switching of patient categories increase hospital revenues. The results of this study suggest that treatment plan switching differs across hierarchical prescribing practices. The results showed that prescribing decisions were based on a unique characteristic of the payment system's prescribing frequency without affecting healthcare costs. In addition, the results suggest that changing treatment plans will improve cost-effectiveness, and hospitals may benefit from prescribing plans as physicians treat a wider range of patients with diverse symptoms. This study uses health analytics management and cloud computing to understand how different health plans work in Thailand.

7 Limitation and Implication of This Research

A limitation of this study is the availability of the electronic medical records database, which may be difficult in rural areas of developing countries. Electronic medical records capture general data, including general patient information, demographic information, health information, treatment history, contact information, and payment information. Using this data, physicians and practitioners can update their information. Some types of recorded health information and treatment history are restricted, such as sensitive conditions, including treatment protection, so retention rates cannot be retrieved. There are also issues with manually entering data input that has not yet been converted to a digital format. This can lead to difficulties in applying research methods and techniques at other hospitals because local staff lack knowledge of data management. This approach prevents local medical staff from performing the procedure without proper training. In general, the results of empirical studies often need to be replicated or evaluated against other considerations when treatment is individualized. However, this study sampled the actual process using electronic records, which may have methodological limitations that prevent physicians from making clear statements that would lead physicians to revise practice guidelines, such as equal treatment for cost-effectiveness. Clear statements about medical practice might be more likely to be based on systematic procedures or clinical practice guidelines that are based on comprehensive literature reviews. Hence, this study has raised some interesting issues about equity in patient care that can be found in management practice reports that support other claims about the impact on the healthcare delivery process. While this study highlights the importance of defining local healthcare processes in Thailand, the findings shed light on the prescribing patterns of rural healthcare services. The importance of this study is to provide an overview for future research that addresses the intersection of economic models and performance-based social healthcare.

References

1. Giancotti M, Guglielmo A, Mauro M (2017) Efficiency and optimal size of hospitals: results of a systematic search. PLoS ONE 12(3):e0174533
2. Snyder CW, Weinberg JA, McGwin Jr G, Melton SM, George RL, Reiff DA, Kerby JD (2009) The relationship of blood product ratio to mortality: survival benefit or survival bias?. J Trauma Acute Care Surg 66(2):358–364
3. Teplensky JD, Pauly MV, Kimberly JR, Hillman AL, Schwartz JS (1995) Hospital adoption of medical technology: an empirical test of alternative models. Health Serv Res 30(3):437

4. Bramlage P, Messer C, Bitterlich NA, Pohlmann C, Cuneo A, Stammwitz E, Tebbe U (2010) The effect of optimal medical therapy on 1-year mortality after acute myocardial infarction. Heart 96(8):604–609
5. Dagenais GR, Lu J, Faxon DP, Kent K, Lago RM, Lezama C (2011) Bypass angioplasty revascularization investigation 2 diabetes (BARI 2D) study group. Effects of optimal medical treatment with or without coronary revascularization on angina and subsequent revascularizations in patients with type 2 diabetes mellitus and stable ischemic heart disease. Circulation 123(14):1492–1500
6. Longford NT (1999) Selection bias and treatment heterogeneity in clinical trials. Stat Med 18(12):1467–1474
7. Mirea L, Sankaran K, Seshia M, Ohlsson A, Allen AC, Aziz K (2012) Canadian neonatal network. Treatment of patent ductus arteriosus and neonatal mortality/morbidities: adjustment for treatment selection bias. J Pediatrics 161(4):689–694
8. Glesby MJM, Hoover DR (1996) Survivor treatment selection bias in observational studies: examples from the AIDS literature. Ann Intern Med 124(11):999–1005
9. Brooks R, Grotz C (2010) Implementation of electronic medical records: how healthcare providers are managing the challenges of going digital. J Bus Econ Res (JBER) 8(6)
10. Ochieng OG, Hosoi R (2005) Factors influencing diffusion of electronic medical records: a case study in three healthcare institutions in Japan. Health Inf Manage 34(4):120–129
11. Sood SP, Nwabueze SN, Mbarika VW, Prakash N, Chatterjee S, Ray P, Mishra S (2008) Electronic medical records: a review comparing the challenges in developed and developing countries. In: Proceedings of the 41st annual Hawaii international conference on system sciences (HICSS 2008). IEEE, pp 248–248
12. Lin HL, Wu DC, Cheng SM, Chen CJ, Wang MC, Cheng CA (2020) Association between electronic medical records and healthcare quality. Medicine 99(31)
13. Abdekhoda M, Dehnad A, Zarei J (2019) Determinant factors in applying electronic medical records in healthcare. East Mediterr Health J 25(1):24–33
14. Bardach SH, Real K, Bardach DR (2017) Perspectives of healthcare practitioners: an exploration of interprofessional communication using electronic medical records. J Interprof Care 31(3):300–306
15. Winkelman WJ, Leonard KJ (2004) Overcoming structural constraints to patient utilization of electronic medical records: a critical review and proposal for an evaluation framework. J Am Med Inform Assoc 11(2):151–161
16. Holroyd-Leduc JM, Lorenzetti D, Straus SE, Sykes L, Quan H (2011) The impact of the electronic medical record on structure, process, and outcomes within primary care: a systematic review of the evidence. J Am Med Inform Assoc 18(6):732–737
17. Johnson SB, Bakken S, Dine D, Hyun S, Mendonça E, Morrison F, Stetson P (2008) An electronic health record based on structured narrative. J Am Med Inf Assoc 15(1) 54–64
18. Marcos M, Maldonado JA, Martínez-Salvador B, Boscá D, Robles M (2013) Interoperability of clinical decision-support systems and electronic health records using archetypes: a case study in clinical trial eligibility. J Biomed Inform 46(4):676–689
19. Underhill K (2013) Study designs for identifying risk compensation behavior among users of biomedical HIV prevention technologies: balancing methodological rigor and research ethics. Soc Sci Med 94:115–123
20. Farber NJ, Cederquist L, Devereaux M, Brown E (2011) Do extratherapeutic factors affect residents' decisions to prescribe medication for erectile dysfunction in ethically challenging scenarios? Acad Med 86(12):1525–1531
21. Weber EU, Blais AR, Betz NE (2002) A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. J Behav Decis Mak 15(4):263–290
22. Chen CHC, Taylor M (2016) An assessment of government regulation on adaptive capability and managerial strategy in US healthcare. Int Manage Rev 12(2):5–19
23. Nunes R, Rego G, Brandão C (2009) Healthcare regulation as a tool for public accountability. Med Health Care Philos 12(3):257–264

24. Hunter BM, Murray SF, Marathe S, Chakravarthi I (2022) Decentred regulation: the case of private healthcare in India. World Dev 155:105889
25. Ellis RP, McGuire TG (1986) Provider behavior under prospective reimbursement: cost sharing and supply. J Health Econ 5(2):129–151
26. Liu YM, Yang YHK, Hsieh CR (2009) Financial incentives and physicians' prescription decisions on the choice between brand-name and generic drugs: evidence from Taiwan. J Health Econ 28(2):341–349
27. Nguyen H (2011) The principal-agent problems in health care: evidence from prescribing patterns of private providers in Vietnam. Health Policy Plann 26(suppl_1):i53–i62
28. Iizuka T (2012) Physician agency and adoption of generic pharmaceuticals. Am Econ Rev 102(6):2826–2858
29. Aitken ML, Berndt ER, Bosworth B, Cockburn IM, Frank R, Kleinrock M, Shapiro BT (2013) The regulation of prescription drug competition and market responses: patterns in prices and sales following loss of exclusivity. In: Measuring and modeling health care costs. University of Chicago Press, pp 243–271
30. Abbott TA III (1995) Price regulation in the pharmaceutical industry: prescription or placebo? J Health Econ 14(5):551–565
31. Kanavos P, Vandoros S (2010) Competition in prescription drug markets: is parallel trade the answer? Manag Decis Econ 31(5):325–338
32. Stein P, Valery E (2004) Competition: an antidote to the high price of prescription drugs. Health Aff 23(4):151–158
33. Huskamp HA, Rosenthal MB, Frank RG, Newhouse JP (2000) The medicare prescription drug benefit: How will the game be played? A nuts-and-bolts proposal for using competition and pharmacy benefit managers to contain drug costs and promote quality. Health Aff 19(2):8–23
34. Farley PJ (1986) Theories of the price and quantity of physician services: a synthesis and critique. J Health Econ 5(4):315–333
35. Bala R, Bhardwaj P, Chen Y (2013) Offering pharmaceutical samples: the role of physician learning and patient payment ability. Mark Sci 32(3):522–527
36. Makoul G, Arntson P, Schofield T (1995) Health promotion in primary care: physician-patient communication and decision making about prescription medications. Soc Sci Med 41(9):1241–1254
37. Barros PP (2011) The simple economics of risk-sharing agreements between the NHS and the pharmaceutical industry. Health Econ 20(4):461–470
38. Constantatos C, Perrakis S (1997) Vertical differentiation: entry and market coverage with multiproduct firms. Int J Ind Organ 16(1):81–103
39. Noh YH, Moschini G (2006) Vertical product differentiation, entry-deterrence strategies, and entry qualities. Rev Ind Organ 29(3):227–252
40. Atal JP, Cuesta JI, Sæthre M (2018) Quality regulation and competition: Evidence from pharmaceutical markets. NHH Department of Economics Discussion Paper, (20)
41. Zakumumpa H, Rujumba J, Kwiringira J, Katureebe C, Spicer N (2020) Understanding implementation barriers in the national scale-up of differentiated ART delivery in Uganda. BMC Health Serv Res 20(1):1–16
42. Adbi A, Bhaskarabhatla A, Chatterjee C (2020) Stakeholder orientation and market impact: evidence from India. J Bus Ethics 161(2):479–496

Knowledge Graph Generation from Model Images

Srinivasan Kandhasamy, Chikkamath Manjunath, Praveen C. V. Raghava, Sandeep Kumar Erudiyanathan, and Gohad Atul Anil

Abstract Model-based system engineering (MBSE) for automotive requirements is gaining prominence over the last decade in the software industry. MBSE models can be designed and constructed for the corresponding unit level requirements to system level requirements using tools like MATLAB/Simulink, ASCET, etc. Once these models are made available, extracting and inferring information from these models will help in understanding on how the corresponding requirements are realized. Here, we tried to develop a technique to capture these model-related data from image files and converting them into equivalent knowledge. The knowledge is extracted, interpreted, and inferred for a given context and stored it in the form of a knowledge graph (KG). There are three major components: element detection, connection analysis, and text analysis. Here each of these components is further explained in detail. The information obtained out of these components is used to generate knowledge graphs. These KGs are further used for various applications in the V-model software development cycle. Test case generation is one among various use cases of a KG. Here brief explanation is provided on an use case where KG is used as a means to create test intents through querying.

Keywords Knowledge graph · Image processing · MBSE · CCA · Text analysis

S. Kandhasamy (✉) · C. Manjunath · P. C. V. Raghava · S. K. Erudiyanathan · G. A. Anil
Bosch Global Software Technologies, Bangalore 560 095, India
e-mail: srinivasan.ks@in.bosch.com
URL: https://www.bosch-softwaretechnologies.com/en/index.html

C. Manjunath
e-mail: manjunath.chikkamath@in.bosch.com

P. C. V. Raghava
e-mail: praveen.cvr@in.bosch.com

S. K. Erudiyanathan
e-mail: erudiya.SandeepKumar@in.bosch.com

G. A. Anil
e-mail: atulanil.gohad@in.bosch.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_3

1 Introduction

In automotive industries, majority of the software applications are developed by either directly through the embedded C programming or modeled graphically using ASCET or MATLAB/Simulink tool. Many static software analysis tools exist for C or Java [3], whereas there is lack of availability of such tools for model-centric development approach like ASCET or MATLAB. These tools are used for developing application software for embedded systems using graphical models and textual programming notations and provide a model-based innovative solution and representation. In model-based development, an executable of the system is generated while establishing its properties through simulation and testing in early stages of development. When the model behaves as required, it can be converted automatically to production quality code. Prior to deployment, testing of these models plays a critical role in avoiding the malfunction of units/false behaviors.

The system or unit models from ASCET or MATLAB/Simulink are complex to understand when the intended applications for such models consist of multiple functional components. These models have complex interactions of many sub-models and components. These sub-model components are connected with different control or signal lines in multiple sequences to perform different actions. The system model eventually calculates required output for the set of given input. These models are deployed in the electronic control unit (ECU) to perform various activities. The automated means of extraction of information from models construct knowledge graph (KG) and represent it in the form of machine-readable format which has numerous applications industry wide. One among them can be seen in [1], where authors discuss the top-down and bottom-up approach of KG creation from structured and semi-structured data. A similar work can be seen in [2]. In the current manuscript, our approach uses image analysis [4] to transform the model images. Here conversion of such images to knowledge and these information are further integrated to generate a knowledge graph is discussed.

2 Image Processing Steps

Three major components are used for KG creation: element detection: where the algorithm detects the control elements and other kinds of blocks in an image; connection analysis: where the algorithm detects and extracts the context of which block is connected which others; text analysis: where the text information of the image is extracted. Each of these components is further explained in detail considering ASCET model in the following sections. The steps and KG generation process are same for MATLAB/Simulink or any other such models.

2.1 Element Analysis and Layer Creation

Element analysis includes primarily, detection of various ASCET elements in the model images. Object recognition is used to detect elements in the model images. Fundamentally, the recognition task is carried out by analyzing the shape of elements in the diagrams [5, 6]. One typical model image is shown in Fig. 1. ASCET golden element image library is created with all the ASCET model icon images from the ASCET installation. Element detection algorithms are used to detect the standard ASCET elements in each of the model images (Fig. 2). Scalar elements are identified by analyzing the shape of the scalar elements using Laplacian convolution operator. Similarly, system elements corners are marked and boundary identification algorithm is deployed to detect the location of the system elements. The scalar elements boundary is reduced in all the sides to mark the scalar text for input and output elements.

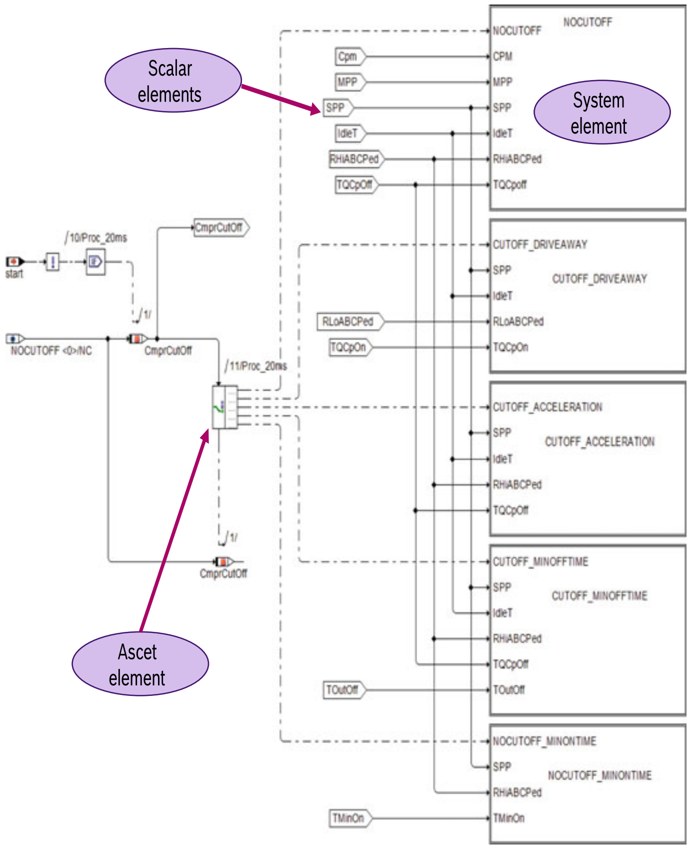


Fig. 1 ASCET Image

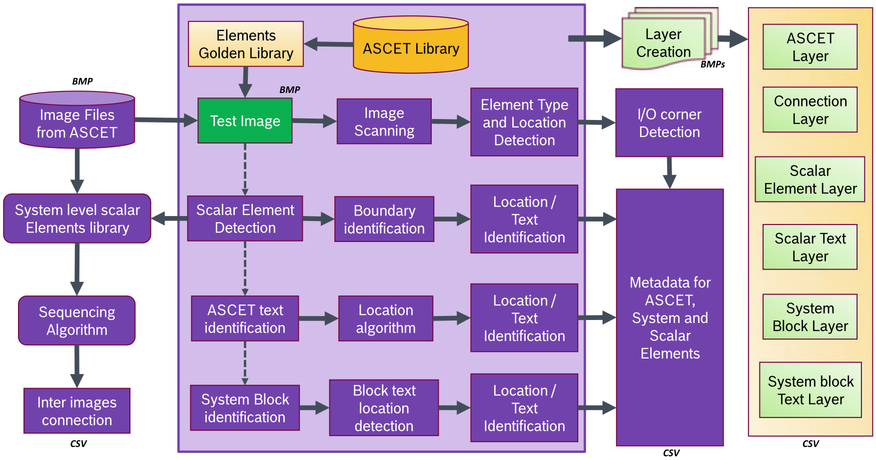
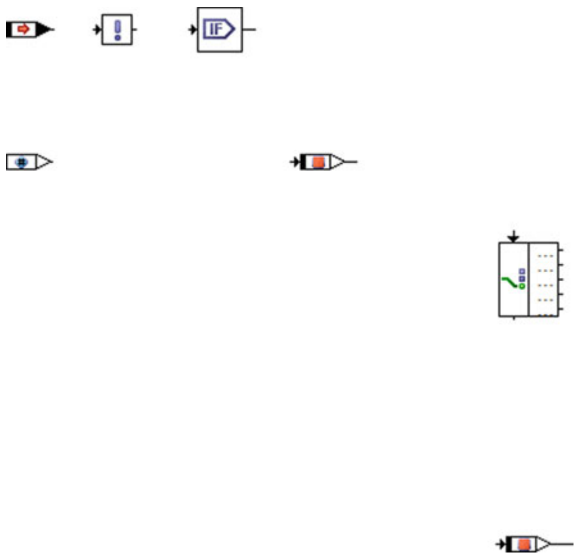


Fig. 2 Element detection and layer creation flow

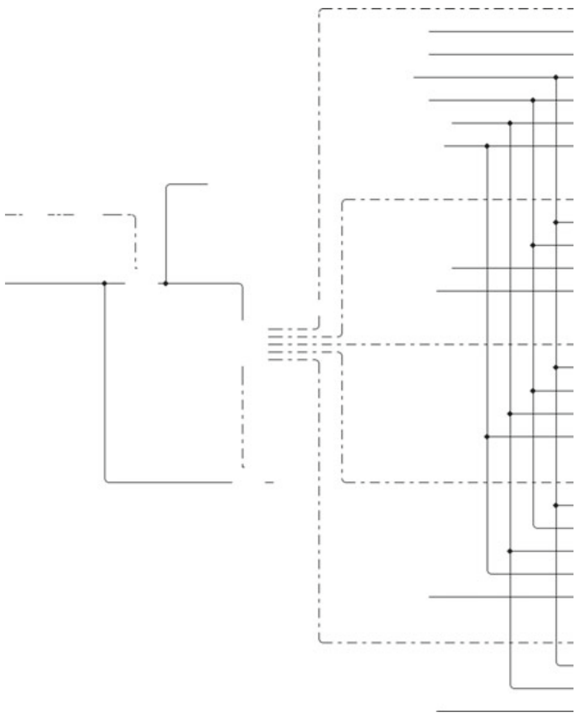
Fig. 3 ASCET Element Layer



Different layers area is generated for further analysis in the pipeline. These layers are created by masking the elements logically and performing arithmetic operations on the images. Finally, the ASCET elements texts are separated from the connections by deploying novel algorithm. This helps in creating connection layer and text layers for further analysis (Figs. 3, 4 and 5).

Scalar elements are the elements connecting all the model images in the image container. These scalar elements from all the images are added in the ASCET template library, and further matching of the scalar elements is performed by further

Fig. 4 Connection layer



detection of this library elements across all the images. This scalar elements identification related to each image and matching them helps in understanding the sequence of the model element connections across images. The connection layer is used as input image for the connection analysis, and scalar element and block element text layers are used as input images for text analysis. In addition to the above output images, element detection is also providing two CSV files. One CSV file contains the information about the location of the element blocks (ASCET, scalar, sequence, and system) and the CG locations. Also it provides the information about the sides if it is input side or output side of the element which is marked as 1 for input side and 2 for output side. This CSV has also the connection point (X,Y) coordinate information.

Harris corner detection algorithm is used for the calculation of the corners for the I/O points of the blocks. The other CSV file contains the information about the sequence elements (scalar) present in the given images. These sequence connections are listed as source and destination elements though multiple elements which are connected in sequence to aid the knowledge graph creation.

Fig. 5 Text layer creation

2.2 Connection Analysis

Connection analysis of a block diagram plays a vital role in extracting the information regarding inter-block dependencies. However, the techniques developed in this project do not impose constraints on the software used, and the experiments are conducted on the ASCET block diagrams. Sophisticated image processing algorithms are used for the connection analysis. The following steps are carried out to perform this task:

1. Input the connection layer image discussed in Sect. 2.1 and the original image to the algorithm.
2. Read the bounding box information of the control elements that were masked in the connection layer/recognized in the element detection phase of the project.
3. Extract the connections and find the bounding boxes. From these connections, extract the corner points. Using the extreme corner points, find the association of the connection to the elements.
4. These elements are given unique IDs, and the end of every connection signifies the mapping of the connection to the block (Figs. 6 and 7).

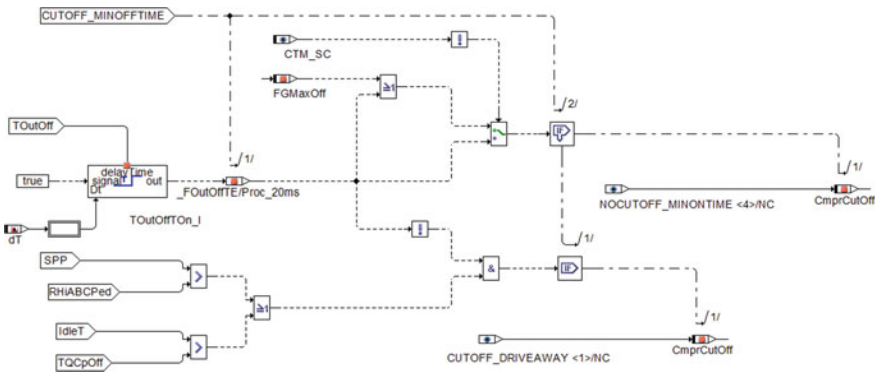


Fig. 6 Image for connection analysis

5. Together with the IDs also the tap points, corner points which are having 90° bend points and association of lines to blocks and line crossing points are identified.
6. Further these points are sorted to extract the context information of which block to which other blocks in an image. Resolving these points plays a critical role in getting the information about block interconnections in an image. In Fig. 7, the tap points are shown in figure with the text “tap”.
7. Further, there are three types of lines in an image: solid, dotted, and sequence lines. Solid lines represent the analog signals flowing from one element to another, dotted lines carry digital signal, and a sequence line carries information about the sequence in which certain elements have to be executed in an image. This classification is carried using a smudge of image processing and deep learning classification algorithms.
8. The obtained image information is later used to chain in a sequence or in the order in which they are executed and transformed into KG using suitable ontological definitions.
9. The overall block diagram representation of the connection analysis is shown in Fig. 8.

2.3 *Text Recognition*

Text recognition is the process of identifying the image characters as part of pre-trained alphabets or symbol. Text identification performs the task of identifying the candidate text region using morphological operations. It does not understand or recognize the precise text in the selected contours. It is imperative we differentiate between on text and non-text characters; this differentiation happens both in text detection and text recognition phases. Text recognition is achieved by the deep learning algorithms. Here in this project we have used the Tesseract Python library to

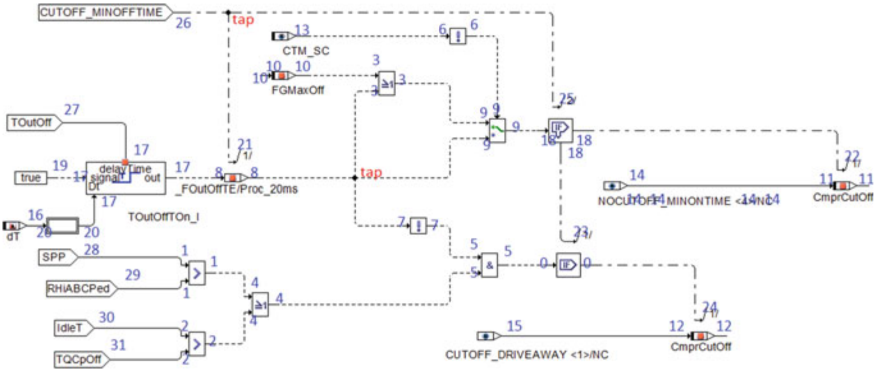


Fig. 7 Identified connections

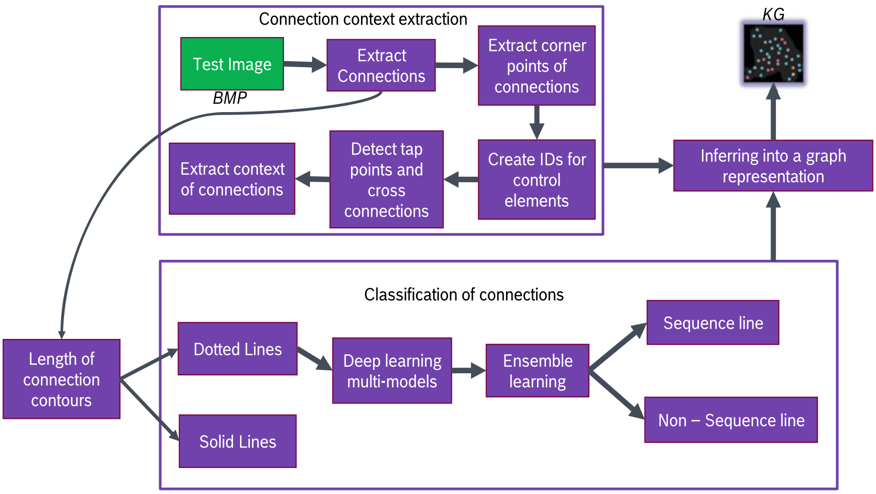


Fig. 8 Connection identification flow

extract the ASCET text from the images. The text recognition is shown in Fig. 9. Text recognition architecture uses concepts of adaptive thresholding, connected component analysis, deep learning methods for text differentiation with other aspects of image, and deep learning methods for word classification and final word output.

3 Knowledge Graph Generation

3.1 Ontological Data Creation

An ontology represents the fundamental knowledge pertinent to the application domain, namely the concepts constituting the domain and the relationships between

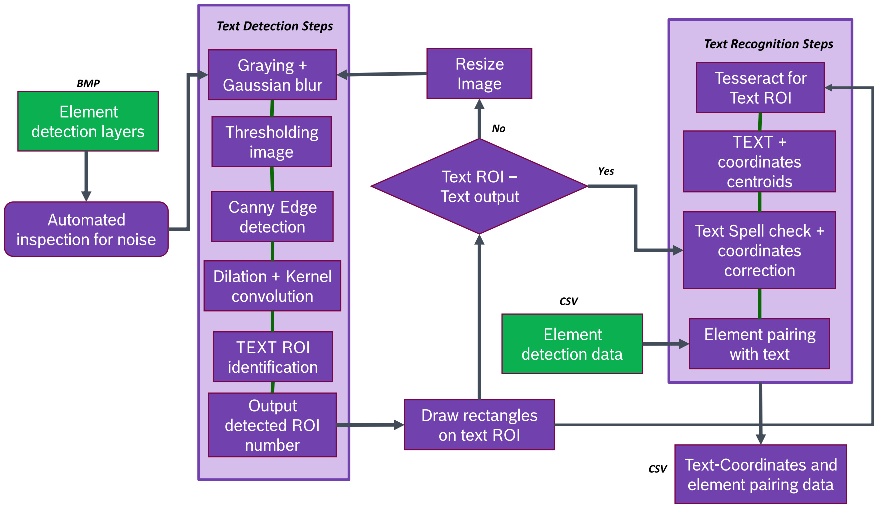


Fig. 9 Text recognition

them. Ontology is semantic data models that define the types of things that exist in any domain and the properties that can be used to describe them. Ontology is generally regarded as smaller collections of assertions that are hand-curated, usually for solving a domain-specific problem. Ontology for ASCET elements has been manually created, using ASCET Automotive System Library and ASCET Icon Reference Guide. This data is kept as a standard library and used in conjunction with the each functional component data fetched from the components discussed in Sect. 2.

3.2 *Knowledge Graph Creation*

Knowledge graph (KG) is a structured form to capture these relationships and entities from various structured and unstructured data sources. ASCET model images containing several model information and related ASCET functions with interdependencies can leverage a KG representation. The quality of KG depends on the nodes and relationships extracted from the data. KG development typically consists of two major phases: knowledge extraction and knowledge completion. Knowledge extraction consists of the tasks NER and RE [7]. These NERs are already available in the form of block names extracted from the model images. The ASCET functions in connection with these blocks provide the necessary information for relation among the nodes (RE). The generated ontology data is imported to Neo4j tool to generate the knowledge graph.

3.3 KG Use Case

In an use case, AI tool is used to generate structured signal-level test specification given software requirements. The idea is to convert non-structured software requirements into structured information and to be able to use it for any automation purpose. It is a pipeline solution with a list of AI/NLP algorithms trained with specific domain corpus [7]. Test case generator is combined with the model KG generated from the model images to automatically generate test cases for given requirements.

4 Conclusion

With the available image repository belonging to a functional component document, the pipeline extracts element information with its boundary point details in an image, connection information details, and text details. The obtained information is used to generate a knowledge graph. A use case to generate the test cases from the requirement to test case is successfully implemented . Under weak constraints, the algorithm works with good accuracy, where the test cases generated suffice the functional requirement of the indented design of the model. Current focus is on improving the overall accuracy and performance of the pipeline to meet various use cases using advanced machine learning and deep learning algorithms and parallel processing wherever applicable.

References

1. Zhao Zhanfang, Han Sung-Kook, So In-Mi (2018) Architecture of knowledge graph construction techniques. Int J Pure Appl Math 118(19):1869–1883
2. Cotter M, Hadjimichael M, Markina-Khusid A, York B (2022) Automated detection of architecture patterns in MBSE models. In: Madni AM, Boehm B, Erwin D, Moghaddam M, Sievers M, Wheaton M (eds) Recent trends and advances in model based systems engineering. Springer, Cham. https://doi.org/10.1007/978-3-030-82083-1_8
3. Klocwork: Best static code analyzer for developer productivity, SAST, and DevOps/DevSecOps https://www.perforce.com/products/klocwork. Accessed 20 Oct 2022
4. Pratt William K (2002) Digital image processing PIKS Inside, 3rd edn. Wiley-Interscience Publication, Wiley
5. Belongie S, Puzicha J (2002) Shape matching and object recognition using shape contexts. IEEE Trans PAMI 24:509–522
6. Gellaboina MK, Venkoparao VG (2009) Graphic symbol recognition using auto associative neural network model. In: Seventh international conference on advances in pattern recognition
7. Veera P, Prasad PVRD, Chikkamath M, Ponnalagu K, Mandadi S, Praveen CVR (2018) Req2Test - graph driven test case generation for domain specific requirement. Int J Comput Trends Technol 60:123–132. https://doi.org/10.14445/22312803/IJCTT-V60P120

Measuring the Performance of An Object-Based Multi-cloud Data Lake

Miguel Zenon Nicanor L. Saavedra and William Emmanuel S. Yu

Abstract As the amount of data generated by society continues to become less structured and larger in size, more and more organizations are implementing data lakes in the public cloud to store, process, and analyze this data. However, concerns over the availability of this data as well as the potential of vendor lock-in lead more users to adopt the multi-cloud approach. This study investigates the viability of this approach in data lake use cases. Results that a multi-cloud data lake can potentially be implemented with less than 1% performance impact to query run times at the cost of a 300% increase in one-time loading. This opens the door for future work on more algorithms and implementations that leverage multi-cloud deployments to enhance availability, scalability, and cost optimization.

Keywords Cloud · Data lake · Data analytics · Big data

1 Introduction

The amount of data being generated today continues to grow at a fast pace. From 2020–2021, the amount of data captured and consumed was estimated to have grown by over 20% from 64.2 zettabytes to 79.0 zettabytes [14]. While the amount of data continues to grow in size, over 80% of this data is considered to be unstructured or semi-structured [3]. Traditional data management and data storage systems such as relational database management systems and data warehouses do not analyze these types of data well as they are primarily designed for structured data. In recent years, data lakes are becoming the more popular approach to analyze and process data because of their ability to handle unstructured and semi-structured data, as well as because of their relatively lower costs when compared to warehouses [10].

M. Z. N. L. Saavedra (✉) · W. E. S. Yu
Ateneo de Manila University, Quezon City, Philippines
e-mail: msaavedra@ateneo.edu

W. E. S. Yu
e-mail: wyu@ateneo.edu

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_4

While various organizations and companies today offer data lake technologies, the most prevalent method of implementing a data lake is through public cloud infrastructure and cloud service providers (CSPs) [2]. The pay-as-you-go billing model as well as the decoupling of storage and compute make the cloud an ideal place for both data lake storage and data lake processing.

As cloud technologies mature, analysts and engineers are beginning to show concern regarding the potential issues and risks of cloud [4]. First, with regard to disaster recovery, if a single cloud provider experiences downtime issues, this could potentially lead to the loss of a large chunk of company data. Secondly, authors have shown concern over the possibility of vendor lock-in [4]. Having multiple application programming interfaces (APIs) with a lack of interoperability with other cloud providers may lead companies to become too dependent on their own provider. Finally, with regards to elasticity, CPU manufacturing shortages have impacted even large cloud providers. There will be great benefit in leveraging multiple CSPs to ensure the availability of processing power.

The goal of this study is to implement a multi-cloud data lake and benchmark its performance. By analyzing the performance impact of spreading data across multiple CSPs, this paper aims to help organizations determine the potential advantages or disadvantages of a multi-cloud data lake solution.

2 Related Works

The trend of data lakes today is to separate compute and storage [5]. This allows each to scale independently of the other and even leverage ephemeral computing environments for batch jobs. Section 2.1 starts by giving an overview of object storage and its overall architecture, while Sect. 2.2 talks about processing mechanisms that run on top of object stores. Finally, Sect. 2.3 discusses the current efforts in building multi-cloud storage and analytic systems.

2.1 Cloud Object Storage

An object represents a file and metadata associated with the file [5]. This includes access control, encryption, and arbitrary user-defined metadata. Objects are immutable constructs that cannot be modified in place but can be deleted and overwritten. These objects are then placed into hierarchical namespaces called buckets. This makes them slow to write but ideal for write-once-read-many (WORM) use cases. The majority of CSPs today implement their own version of object storage. These include the Amazon Simple Storage Service (S3), Azure Blob Store, and Google Cloud Storage (GCS). They offer "unlimited" amounts of total storage, the ability to seamlessly scale based on the number of storage clients, and a pay-as-you-go pricing with no commitment requirement.

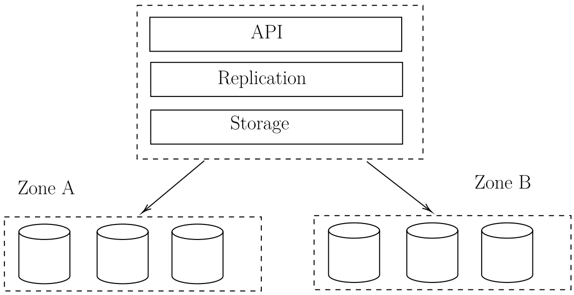


Fig. 1 Components of cloud object storage

Fig. 1 illustrates the general architecture of object stores implemented by most cloud providers. The API layer accepts the raw file data as well as any metadata input. This data is then sent to the replication layer and duplicated across multiple data centers and disaster zones[1] within the same geographic area.

Object stores are an ideal system for data lake storage as lakes are primarily read-heavy workloads. Data is collected and ingested from different data sources and stored in an object storage system [15]. A separate system is then used to catalog additional metadata for the data lake such as partition information and table schema. Finally, processing systems leverage the data catalog for schema and location references to run queries on top of the object store without the need to ingest any data.

2.2 Distributed Processing

To ensure that the storage and compute components are decoupled, distributed computing environments are normally provisioned separately from the object stores, then used to query the data from the lake [5]. This type of processing became more common after the introduction of the Hadoop ecosystem. Hadoop acts as the processing layer which uses the MapReduce framework to easily distribute query workloads across multiple machines. This is further enhanced by the use of common interfaces such as HiveQL, Hadoop's SQL interface to MapReduce.

As the cost of hardware continued to decrease, processing was offloaded from disk into memory. Tools like Spark and Presto can query data from a data lake while keeping the data in-memory for faster processing [6]. This led to query jobs that were 10–100 times faster than plain Hadoop and MapReduce.

The combination of in-memory processing provided by modern frameworks, reliable and scalable object storage, and ephemeral compute environments provided by CSPs all dramatically lower the costs of managing a data lake without sacrificing performance. It has become common practice to keep data stored in object storage but keep processing nodes offline [5]. Because CSPs do not require computing commitments, and nodes are only billed when they are running, users are only billed for storage costs, allowing them to save on processing costs. This setup is often referred to as an *ephemeral* computing environment.

[1] Often known as availability zones or simply, zones.

2.3 Multi-cloud Projects

The use of multi-cloud technologies can further improve the resilience of data lakes and potentially reduce costs by leveraging various pricing models and avoiding vendor lock-in [4]. While research in multi-cloud systems is a growing field, most current work is focused on optimizing for compute, storage cost, availability, and network latency of basic storage operations.

The *Scalia* project was developed to support multi-cloud architectures and improve availability using erasure coding to distribute objects in chunks and spread them across CSPs [8]. They provide an Amazon S3-compatible API as the front-end of their system while the back end handles striping and uploading chunks to multiple supported cloud providers.

MinIO focuses more on scaling cloud object storage by making use of containerized environments. It uses Kubernetes, a ubiquitous container control plane, to distribute the storage workload across multiple containers hosted on virtual machines running on different CSPs [9]. The focus was less on using the native object storage, but more on leveraging the compute resources available for better flexibility and system control.

Noobaa is another multi-cloud project owned by RedHat that acts as a storage gateway to multiple CSPs [11]. Similar to Scalia, they also implement erasure coding to spread object chunks across cloud providers and also implement an S3-based API front-end. Noobaa, however, is deployed in a containerized cluster, allowing it to scale horizontally to handle more traffic CSPs while appearing to clients as a single unified file system.

One point of interest across these studies is that, like more object storage systems today, they have chosen to implement the S3 API over their own custom API. This goes beyond multi-cloud systems as several on-premise and even CSPs[2] have also been implementing the S3 API as a front-end to their storage environments [1]. This is likely due to the support many connectors and tools have for this S3 API.

All these multi-cloud projects focus primarily on handling basic storage operations such as reads and writes. Current studies have not yet explored the use case of analytics workloads in multi-cloud systems which this work hopes to evaluate.

3 Methodology

This section discusses the multi-cloud system used as a data lake as well as the experimental design used to test its performance.

[2] Google Cloud Storage, one of Amazon S3's competitors, has even chosen to implement it.

3.1 System Overview

The system used for this study consists of four major parts as seen in Figs. 2a–b: a multi-cloud storage, a Hadoop cluster for data analytics, Amazon S3 Storage, and Google's GCS. To ensure consistency, all components were hosted in the Singapore region of their respective CSPs.

Storage Gateway This study made use of the Noobaa multi-cloud storage gateway. The storage gateway consisted of an operator which processes user input, a core component that handles storage operations, and a database (DB) component that holds storage metadata. The system was deployed in a Kubernetes (K8s) cluster with three worker nodes as seen in Figs. 2a–d.

The cluster was also deployed on AWS. Each node had 2 virtual CPU cores (vCPUs) and 8 GiB of memory. An elastic load balancer (ELB) was also used to distribute the traffic to the appropriate services. Finally, the S3 API was used as the gateway's API front-end.

Object Storage Two object storage services were used for this study. Amazon S3 was used as it is in AWS with all the other infrastrutcure, while GCS was the external object store hosted in a different on Google Cloud Platform.

Hadoop Cluster The Hadoop cluster in this study was deployed on the Amazon Elastic MapReduce Service (EMR). It consisted of one master node with 4 vCPUs and 16 GiB of memory and four worker nodes, each with 4 vCPUs and 32 GiB of memory. The following Hadoop components were then installed in the cluster: The Hadoop Filesystem (HDFS), Yet Another Resource Negotiator (YARN), Hive, and Presto.

HDFS and *YARN* are standard members of the Hadoop ecosystem that are installed in all Hadoop distributions by default [16]. HDFS acts as the distributed storage layer of Hadoop. While it was not used to store any data for this experiment, it was used to intermediate results from the queries. YARN is Hadoop's resource manager and distributes cluster resources to running jobs.

Hive was the software used as the data catalog for this system. It held the metadata for the data lake, including the locations of the objects being queried, table definitions and schema, and partition information. The Hive s3a connector was also used to connect to the storage gateway.

Finally, *Presto* was used as the query engine for tests. Presto connects to the Hive data catalog to identify the structure of the data then streams data from the object store to analyze the data in-memory [13]. The processing is distributed across the nodes, allowing Presto to scale linearly.

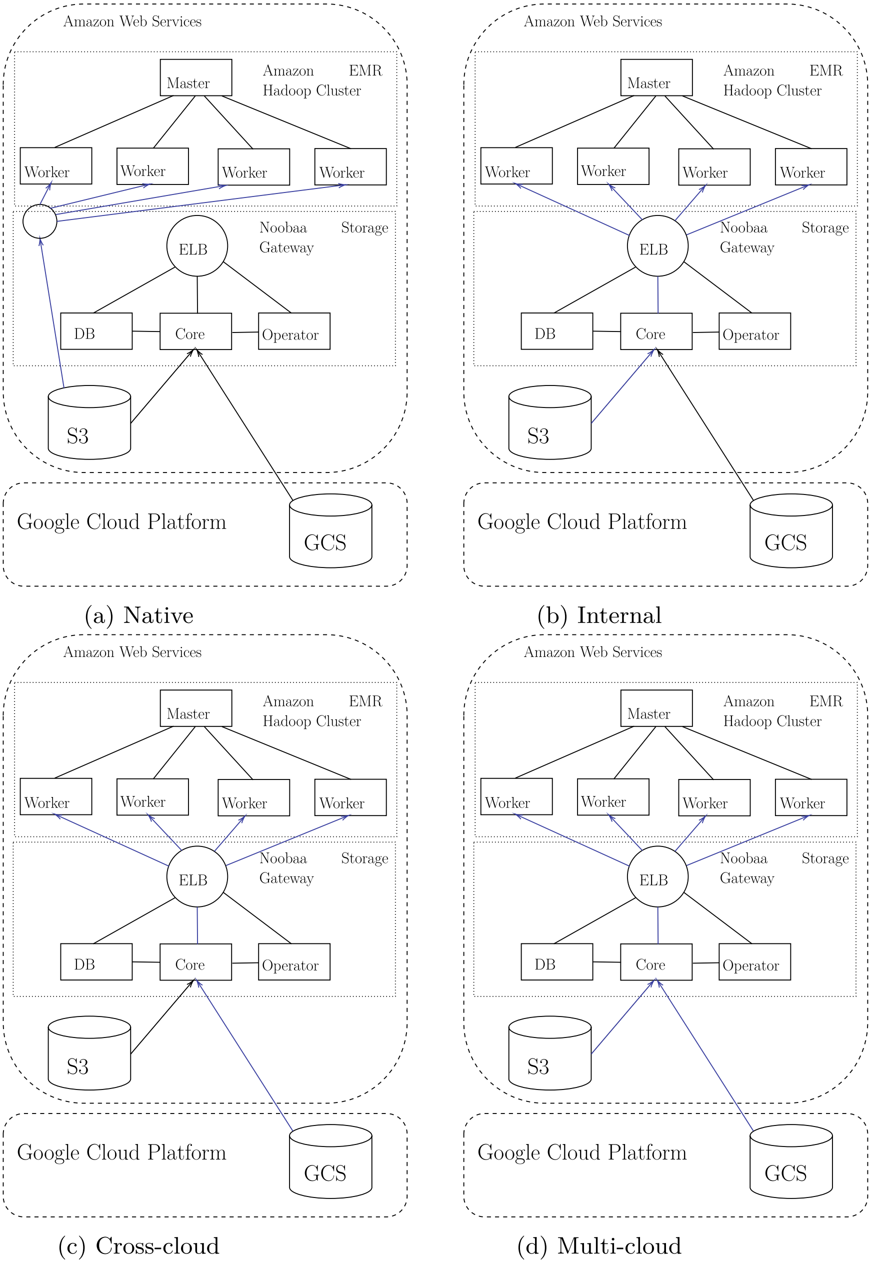


Fig. 2 Data distribution methods

3.2 Data Loading

The New York City Taxi and Limousine Commission (NYC TLC) Trip Record data [7] was used as the sample dataset for this study. The main table consists of 1,547,741,381 rows and 24 columns stored in Parquet format. The data was then loaded into Amazon S3 and partitioned by year and month. To facilitate join queries, another table was also created with 265 rows and 4 columns. This is used as a lookup table for the locationid column.

This data was loaded from a single virtual machine hosted in AWS and was distributed using four methods illustrated in Figs. 2a–d; the blue lines highlighting where the data is passing for each method. The loads were also run 30 times for each method to ensure statistical significance.

Native The first method uses a CSP's native object storage, Amazon S3, to store the data. This is the current method of data lake implementation and serves as the control group and baseline for this comparison [15].

Internal This second method uses the storage gateway but keeps all the data inside the same CSP and therefore the same cloud network. This is to help identify if the impact of the overhead from the storage gateway. This method made use of infrastructure and services that were solely on AWS.

Cross-cloud The third method stores all the data completely in the other CSP. This represents the case where the processing cluster will need to retrieve all the data from another CSP and therefore another cloud network. This method made use of infrastructure that was solely on AWS, but data stored only in GCS.

Multi-cloud The final method evenly spreads the data across two CSPs. This represents a true multi-cloud scenario where two object stores are in use at the same time. This method made use of infrastructure that was solely on AWS, but half of the data in GCS and half in S3.

3.3 Querying the Data

There were four types of queries used for testing[3]

1. A *filter* query that just scans and filters the data based on trip_distance
2. An *aggregate* query that gets the average fare distance by year and vendor
3. A *join* query that identifies trips within the same borough
4. An aggregate and join (*Agg-Join*) query that computes the average fare by year and vendor for trips in the same borough

[3] Exact queries may be found here: https://gist.github.com/zzenonn/de669bb5ea5393ae853a57fb5f13f806.

Each query was run 30 times for each method defined in Sect. 3.2 to ensure a statistically significant result. A two-tailed homoscedastic t-test was then run to compare the different data distributions against the Native method.

4 Results and Analysis

This study uses two main key performance indicators: *data load time* and *query run time*.

4.1 Data Load Time

Fig. 3 illustrates the data load time per distribution style. The storage gateway noticeably adds a 300% increase in data load time. This is likely due to the additional overhead caused by erasure coding and the context switching between CSPs. To improve the availability of the system, objects were further subdivided into parts before being uploaded and spread across cloud object stores.

There is also a slight increase in load time, and the more data is stored in the separate CSP. The internal distribution shows the lowest load times behind the storage gateway likely because there is no need to traverse cloud provider networks. The cross-cloud distribution shows the highest data loading time as the data all need to be uploaded to a separate cloud provider. The multi-cloud distribution sees a slight improvement over the cross-cloud primarily because part of the data stays in the current CSP.

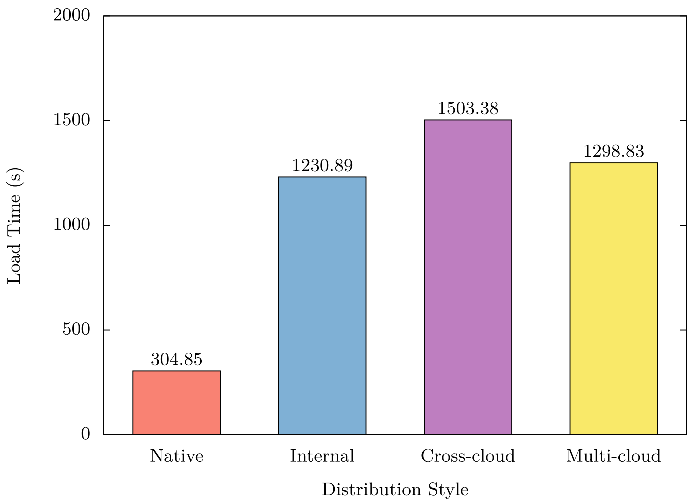


Fig. 3 Average data load time

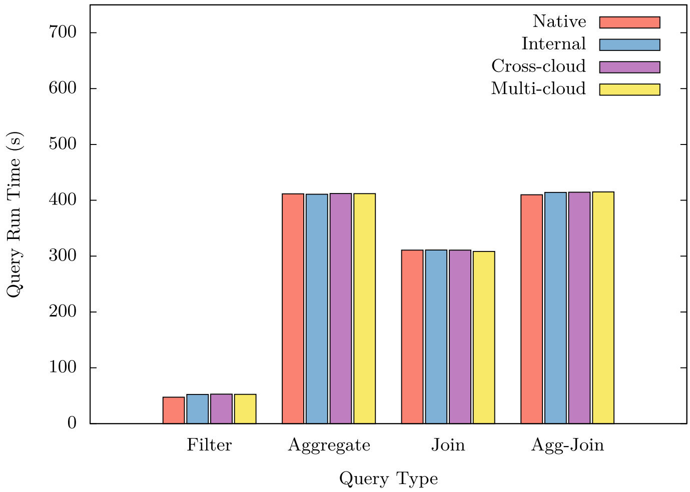


Fig. 4 Average Query Duration

4.2 *Query Run Time*

Fig. 4 shows the average query run time per query and distribution method. The standard deviation for the filter queries ranged from 2 to 3 s across distribution methods, while the other query types' ranged from 7 to 10 s which shows consistent performance in the trials.

When using the multi-cloud gateway, the filter queries show a statistically significant difference when compared to the native method of querying the data lake directly from the object store with $p < 0.01$. The filter queries took 10–11% or approximately 5 s longer to complete when done through the multi-cloud gateway. However, this is likely because filter queries are generally the least complicated and fastest running queries [12]. This means that a even slight variation in execution time has a much larger effect as seen in the results.

For aggregate, join, and aggregate–join queries, t-test results all scored $p > 0.05$, meaning there is no significant difference between the native, internal, cross-cloud, and multi-cloud methods when it comes to these types of queries. There is also only a slight increase not exceeding 1% in average query run time. These results show that the main bottleneck after loading is not the streaming of the data from the object storage but the actual processing of that data.

5 Conclusions and Future Work

Multi-cloud data lakes could potentially increase the availability, durability, and fault-tolerance of large unstructured and semi-structured data. This study evaluates the viability of a multi-cloud solution for data lake use cases. The results show that because of the lack of significant difference in query run times for most query types, multi-cloud data lakes can be implemented without significant impact on query performance.

At the same time, it is important to note that further improvements may be made in the load times of the data sets. Extract, transform, and load (ETL) processes are a very important component in data lakes, and slower load times adversely impact the whole data pipeline as new data may take longer to ingest. However, this load penalty can be seen as a setup time constraint and initial investment for more elasticity and flexibility when the lake is in use. As a single machine was used to perform the load in this study, load times may also be potentially improved with linearly scalable distributed loads.

Future work in this area may look into running tests on datasets of varying sizes to see if the performance impact of the multi-cloud storage gateway scales linearly. Improvements may also be made to the placement algorithms to leverage the varying cost optimization techniques offered by different CSPs by incorporating storage tiering to lower storage and transfer costs and provisioning lower-cost computing environments by optimizing purchase options.

References

1. Dorji U (2018) List of S3 compatible storage providers. https://help.servmask.com/knowledgebase/list-of-s3-compatible-storage-providers. Accessed 16 Aug 2022
2. Grossman RL (2019) Data lakes, clouds, and commons: a review of platforms for analyzing and sharing genomic data. Trends Genet 35(3):223–234
3. Harbert T (2021) Tapping the power of unstructured data. https://mitsloan.mit.edu/ideas-made-to-matter/tapping-power-unstructured-data. Accessed 16 Aug 2022
4. Hong J, Dreibholz T, Schenkel JA, Hu JA (2019) An overview of multi-cloud computing. In: Web, artificial intelligence and network applications. Springer International Publishing, pp 1055–1068
5. Kumar P (2017) Cutting the cord: separating data from compute in your data lake with object storage. https://www.ibm.com/cloud/blog/cutting-cord-separating-data-compute-data-lake-object-storage. Accessed 16 Aug 2022
6. Mami MN, Graux D, Scerri S, Jabeen H, Auer S (2019) Querying data lakes using spark and presto. The world wide web conference. WWW '19. Association for Computing Machinery, New York, NY, USA, pp 3574–3578
7. NYC Taxi and Limousine Commission: TLC trip record data (2022)
8. Papaioannou TG, Bonvin N, Aberer K (2012) Scalia: an adaptive scheme for efficient multi-cloud storage. In: SC '12: Proceedings of the international conference on high performance computing, networking, storage and analysis. IEEE, pp 1–10
9. Pérez-Colado IJ, Pérez-Colado VM, Martínez-Ortiz I, Freire M, Fernández-Manjón B (2020) A scalable architecture for one-stop evaluation of serious games. In: Games and learning alliance. Springer International Publishing, pp 69–78
10. Ravat F, Zhao Y (2019) Data lakes: Trends and perspectives. In: Database and expert systems applications. Springer International Publishing, pp 304–313
11. Red Hat: Red hat OpenShift container storage (2022)
12. Saavedra M, Yu W (2017) A comparison between text, parquet, and pcap formats for use in distributed network flow analysis on hadoop. J Adv Comput Netw 5(2):59–64

13. Singh Y, Kandah F, Zhang W (2011) A secured cost-effective multi-cloud storage in cloud computing. In: 2011 IEEE conference on computer communications workshops (INFOCOM WKSHPS). IEEE, pp 619–624
14. Statista Research Department: Big data - statistics & facts. https://www.statista.com/topics/1464/big-data/ (2022). Accessed 16 Aug 2022
15. Vogels W (2020) How amazon is solving big-data challenges with data lakes. https://siliconangle.com/2020/01/30/amazon-solving-big-data-challenges-data-lakes/. Accessed 16 Aug 2022
16. White T (2015) Hadoop: the definitive guide. "O'Reilly Media, Inc." Google-Books-ID: drbI_aro20oC

A Short Sketch of Solid Algorithms for Feedback Arc Set

Robert Kudelić

Abstract Feedback arc set was presented by Karp in his seminal paper as NP-complete and was being tackled with before and after by various procedures. Number of devised algorithms for the problem thus far is vast, this paper therefore selects a few advantageous algorithms that should generally be favorable and that will typically satisfy readers requirements. This paper delivers synthesis of important elements, for an important problem, in a compact manner, and presents to the reader both algorithms and skeletal information, while at the same time directing to sources of interest for an in-depth purpose.

Keywords Feedback arc set · Algorithms · General application · Skeletal information · In-depth sources

1 Introduction

The problem of feedback arc set (FAS) is quite well-known to the scientific community. It is one of those original problems presented by Karp to be NP-complete by reducing from the problem of Node[1] Cover [12]. The problem has both optimization and decision version, with the decision version being as follows—taken from [8, 14].

Question: "Directed graph $G = (V, A)$, a positive integer $K \leq |A|$." **Answer**: "Is there a subset $A' \subseteq A$ with $|A'| \leq K$ such that A' contains at least one arc from every directed cycle in G?"

The problem is NP-hard [17], APX-hard [11] and in general quite difficult to solve. There are, however, instances that admit polynomial time method. If, for example, one has a graph that is undirected, then a solution can be easily found by minimum spanning tree algorithm [7]. There is also, among others, an instance where input for

[1] MorewidelyknownasVertexCover.

R. Kudelić (✉)
Faculty of Organization and Information Science,Varaždin42000, Croatia
e-mail: robert.kudelic@foi.unizg.hr

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_5

an algorithm is a graph well-known as reducible flow graph. In such a situation, FAS can be efficiently solved in polynomial time [16, 19]—for more details, one should consult [14].

The problem of FAS is practically important and can be often found in various situations. Some of these are machine learning [3], search engine ranking [18], computational biology [9], cryptography [20], etc. Therefore, tackling with FAS has much wider implications than only those theoretical ones.

The problem of FAS is quite old and goes all the way back, even unto 1960s [14]. It has been "attacked" from various sides and in many ways—and as stated by Peter Eades[2]: "The problem is well-known and has been investigated by the best minds in Computer Science". For such a historical, but also state of the art, review, the reader may consult [6, 14]—as such a review would go out of scope of this paper.[3]

It is the aim of this paper to emerge, from such a vast collection of algorithms for FAS, a few procedures that are efficient, solution-wise or guarantee-wise favorable, and preferably not overly complex to implement. Since one typically does not need a procedure that will give better result, but will seriously lack in efficiency, or vice versa, or a procedure that requires strong effort to implement.

Therefore, a reader will be able to quickly find solid algorithms that will be fast for a purpose and that will output quality solutions. By having in his toolbox, the following select algorithms for FAS, one will be widely prepared for tackling with FAS and will not be encumbered with time and effort needed to comb vast number of algorithms.

The algorithms selected for which it has been shown through scientific research excellence in efficiency or approximal quality or solution optimality are GreedyFAS from [5] (originally called GR by Eades et al. in [5], but later on in [21] renamed GreedyFAS), BergerShorFAS from [2] (as well coined in [21]), MonteCarloFAS from [13] (metaheuristic improvement is published in [15]) and ExactFAS from [1]. These algorithms will be presented through important details and with brevity—in such a way goal of this research will be achieved.

2 Algorithm GreedyFAS

The main idea of this algorithm is to iteratively remove graph nodes that are sources and sinks, and those nodes for which $\delta(u)$,[4] where u represents graph node, are currently maximum in value [5].

This series of steps is executed in such a way where all "sink-like" nodes are placed at the end of some ordering π, and all "source-like" nodes are placed at the beginning of the same ordering—with the obvious goal of minimizing feedback arc sum [5,

[2] In the forward of [14].

[3] There is also a companion Website for [14] where one can find additional details: https://cs.foi.hr/fas/book/.

[4] $\delta(u) = \delta^{+}(u) - \delta^{-}(u)$.

Algorithm 1 GreedyFAS [5]

```
s1 ← 0; s2 ← 0
while G ≠ 0 do
    while G contains a sink do
        choose a sink u; s2 ← us2; G ← G − u
    end while
    while G contains a source do
        choose a source u; s1 ← s1u; G ← G − u
    end while
    choose a vertex u for which δ(u) is a maximum
    s1 ← s1u; G ← G − u
end while
return s ← s1s2
```

14]. Pseudocode for this procedure is presented in algorithm 1. As an initialization step to the algorithm a bin (bucket) sort can be executed so as to partition vertices into sources, sinks, and δ-classes [5].

The algorithm is quite fast, and its complexity is linear, $O(m)$, where $m = |A|$. [5] Algorithm solution is bounded, $|R(s)| \leq m/2 - n/6$, where $n = |V|$. [5] This algorithm was best performer in the Web-scale research,[5] results of which can be found in [21].

An array implementation, mimicking list behavior, of GreedyFAS found in [21] is achieving $O(m + n)$. [14, 21] Experimental research has showed that FAS size returned by the algorithm is "drastically smaller than the size suggested by the worst-case bound" [14, 21].

Main paper for this algorithm is [5], with [21] being multi-algorithm comparison research and [14] being comprehensive FAS book.

3 Algorithm BergerShorFAS

The main idea behind this randomized algorithm stems from FAS dual, namely maximum acyclic subgraph (MAS) [2]. The idea is to process graph $G = (V, E)$ vertices according to their in/out-degree [2].

If at each iteration, considering some permutation π and processing vertices successively, arc set of bigger size is chosen and added into E', then the resulting subgraph will be large in terms of number of containing arcs and will also be acyclic [2, 14]. When execution of the algorithm is finished, $G' = (V, E')$ will be an acyclic subgraph, and $E \setminus E'$ will consequently be set of arcs without which G is acyclic, i.e., feedback arc set [2, 14].

Similarly, algorithm for FAS can be and is designed, pseudocode of which is presented in algorithm 2. The adapted algorithm computes feedback arc set F directly—without the need for E', which results in the algorithm that is less demanding in terms of memory [21].

[5] Additional details is given in the discussion section.

Algorithm 2 BergerShorFAS [14, 21]

Input: Directed graph $G = (V, E)$.
Output: A feedback arc set for G.

fix an arbitrary permutation π of the vertices of G
$F \leftarrow 0$
for all vertices v processed in order based on π **do**
  if $inDegree(v) > outDegree(v)$ **then**
    $F \leftarrow F \cup \{(v, u) : u \in G.succ(v)\}$
  else
    $F \leftarrow F \cup \{(u, v) : u \in G.pred(v)\}$
  end if
  $E \leftarrow E \setminus (\{(v, u) : u \in G.succ(v)\} \cup \{(u, v) : u \in G.pred(v)\})$
end for
return F

BergerShorFAS returns a feedback arc set that is "reasonably small" [21], while at the same time having linear time complexity of $O(m + n)$, where n and m represent vertices and arcs, respectively [21]. Experimental evaluation has revealed that the algorithm "far outperformed the worst-case bound" [14] stemming from the MAS version where acyclic subgraph contains at least

$$\left(\frac{1}{2} + \Omega\left(\frac{1}{\sqrt{d_{\max}}}\right)\right)|E| \tag{1}$$

corresponding graph arcs [21].

The main paper for this algorithm is [21], with the work from [2] being a foundation for the FAS version and [14] being FAS monograph.

4 Algorithm MonteCarloFAS

This algorithm is a randomized algorithm, namely Monte Carlo, with the main idea being uniformly choosing and removing multi-graph arcs, and in such a way "guess" an optimal solution with certain probability [13].

Input for the algorithm is multi-graph ("allows multiple arcs between every pair of nodes, has no loops, and has no arc weights" [14]); therefore, if the graph is not in such a form it has to be transformed into one [13].

The algorithm is uniformly breaking arcs until a state is achieved where multi-graph has become acyclic [13, 14]. At this point, the algorithm will find a permutation π, via topological sorting[6] (TS), and output feedback arc set sum together with accompanied probability—it is assumed that input graph has at least one cycle, which can easily be checked via TS algorithm [13, 14].

Since pseudocode for MonteCarloFAS algorithm is not as compact, we will here give a more streamlined and somewhat different version than the one previously

[6] For topological sorting one can check [4, 10].

Algorithm 3 MonteCarloFAS [13]

Input: Multi-graph $G = (V, A)$.
Output: Probability P, sum of weights $\sum_{1 \le u \le (|V|-1)}^{(u+1) \le v \le |V|} W(v, u)$, TS permutation π.

```
while TS on G = (V, A \ {(u, v)_1, (u, v)_2 ...}) not found do
   uniformly pick an arc (u, v)
   determine via DFS if (u, v) is part of a cycle
   break an arc (u, v) belonging to a cycle,
     return to pick an arc otherwise
   for last (u, v) broken memorize |{(u, v)_1, (u, v)_2 ...}|,
     return to pick an arc otherwise
end while
return P, Σ_{(u,v)∈E,u>v} W(u, v), π
```

published, and with TS check built in.[7] Pseudocode in algorithm 3 can be run multiple times and in turn find improved solutions with higher probability for being optimum.

MonteCarloFAS is solving minimum FAS in polynomial time with arbitrary probability [13]. Algorithm complexity was $O(k|V|^3)$, where k represents number of iterations, and V stands for graph nodes [13]. While the probability that afters k iterations the algorithm has returned optimal solution is

$$1 - \left(1 - \left(\frac{n-2}{n}\right)^{\frac{n}{2}}\right)^k \tag{2}$$

where n is number of nodes of a graph. [13] During experimental part of the research, it has been discovered that the algorithm either finds optimal solution, or a solution "that is on average 3% away from optimum" [14].

The main paper for this algorithm is [13], with [15] being a hybrid between Monte Carlo and ACO, while [14] represents FAS monograph.

5 Algorithm ExactFAS

Main idea for this algorithm[8] is enumeration of all simple cycles in a slow manner; in this way, an incomplete cycle matrix is extended iteratively on sparse graphs[9] by formulation of minimum FAS through minimum set cover [1].

The approach is based on a set cover formulation, and if enumeration of all simple cycles in a graph is tractable, the integer program that has cycle matrix completed can

[7] For a summary of details for MonteCarloFAS, one should consult [14], while the complete information is published in [13]—ant colony optimization (ACO) inspired version that has learning mechanism built in can be found in [15].

[8] This approach was not tested on real graphs [1, 21]; the authors have named this approach Integer Programming with Lazy Constraint Generation. [1].

[9] It is possible that sparse graph has $\Omega(2^n)$ simple cycles (these graphs do appear in practice). [1, 14].

Algorithm 4 ExactFAS [1]

Input: Directed graph G with m edges and non-negative edge weights.
Output: A minimum weight FAS – on integer program P.

compute FAS $F^{(0)}$ of G using for example minimum Set Cover heuristic
$F^{(0)}$ is set as best feasible solution $\hat{y}$ to integer program P
calculate first cycle matrix $A^i | i = 1$, for G and $F^{(0)}$
for $i = 1, 2, \ldots$ **do**
  invoke integer programming solver on incomplete problem $\tilde{P}^{(i)}$
  set lower bound **if** new solution is more expansive
  if lower and upper bounds are equal, **return** optimal $\hat{y}$
  if $G^{(i)}$ without current FAS edges can be TS sorted,
    return optimal solution $y^{(i)}$
  calculate $F^{(i)}$ for $G^{(i)}$ using FAS heuristic
  $y^{(i)}$ is now a feasible solution to P
  set upper bound if $y^{(i)}$ is narrower
  $\hat{y} = y^{(i)}$
  calculate $A^{(i+1)}$ with $G^{(i)}$, $F^{(i)}$ and $A^{(i)}$
end for

then be input to integer programming solver [1, 14]. The cycle matrix used for the program can be reduced during presolving by iteratively removing rows and columns that are dominating and dominated, respectively, together with removal of columns "that intersect a row with a single nonzero entry" [1, 14].

Integer program formulation can be found in [1] under "integer programming formulation as minimum set cover", or in [14] under sub-chapter 3.42. Cycle matrix[10] $A = (a_{ij})$ holds information about whether edge j belongs to cycle i [1].

Pseudocode for the ExactFAS can be seen in algorithm 4. Original procedure is not compact; therefore, we will here give a skeletal and a concise version that will be more easily grasped—while for more details, one can consult either the paper itself [1] or FAS monograph [14].

Experimental part of the research has showed that as the input graph becomes denser, so the median execution time for the algorithm follows—algorithm[11] performed efficiently on sparse graphs [1, 14]. Testing was conducted on sparse graphs, random tournaments, complete graphs, etc., and the results varied depending on size and type of a graph [1]. Execution time ranged from approx. 1 second to thousands of seconds [1]. "In cases encountered during research, only a tractable number of cycles had to be enumerated until a MFAS is found" [14].

The main paper for this algorithm is [1], with [14] being FAS monograph where the algorithm forms a part of a much larger picture.

[10] For an exact algorithm one should consult algorithm 2, "extending the cycle matrix given an arbitrary feedback edge set," of the paper itself [1].

[11] Source of the method can be obtained at: https://sdopt-tearing.readthedocs.io/en/latest/; test graphs and results are available as well.

6 Discussion

The paper presents four algorithms in order: GreedyFAS [5], BergerShorFAS [2], MonteCarloFAS [13], and ExactFAS [1]. Together, these algorithms should cover most instances for a practical purpose, while they also give some insight into where the science is at the moment. Each one, however, has more narrow applicability in regard to their own characteristics.

GreedyFAS, in algorithm 1, represents a heuristic algorithm that is straightforward in its idea. The algorithm is very fast and runs in linear time [5]. Solution of the algorithm is bounded [5], and the algorithm is capable of scaling to extra large problems of billions of arcs "while being a fast algorithm in general" [21]. As a heuristic algorithm, optimality is not guarantied, and outside calculated bound quality of the solution is not known, but has been statistically determined [21].

BergerShorFAS, in algorithm 2, represents a randomized heuristic approach with simple but effective idea behind it. Algorithm runs in linear time [21] and has a worst-case bound stemming from maximum acyclic subgraph algorithm [2]. As a heuristic algorithm, optimality is not a guarantee, and outside calculated bound quality of the solution is not known, but has been statistically determined [21]. BergerShorFAS is comparable, although inferior, with GreedyFAS in terms of solution, and superior in terms of running time [21].

MonteCarloFAS, in algorithm 3, represents randomized approximation method with simple idea through which solution quality is ascertained. This algorithm has polynomial complexity [13] and is finding optimal solution with arbitrary probability (the algorithm is run multiple times so as to increase a chance to "guess" optimal solution) [13]. Optimality is not guaranteed but is approximated through probability [13]. MonteCarloFAS cannot scale and tackle within reasonable time inputs as large as GreedyFAS and BergerShorFAS can, nor is its efficiency on par. But it can offer, for a problem that is hard to approximate [11], arbitrary confidence in a solution [13].

ExactFAS, in algorithm 4, represents an exact method for sparse graphs. The method enumerates simple cycles iteratively by extending cycle matrix in steps [1]. Algorithm performs efficiently on sparse graphs, and it is, however, possible for sparse graphs to have $\Omega(2^n)$ simple cycles, thus inhibiting efficiency and establishing intractability [1]. As an exact algorithm, it has its optimality as a guarantee, but at the expense of efficiency [1]. Convergence varies on type of input graph and is built for a particular purpose, namely simple cycles [1]. This algorithm stands as an option when intractability is not an issue.

7 Constraints

Algorithms presented through this paper have never been evaluated in a single research, head-to-head. Considering the results from available literature, it is not likely that such comparison would produce different results, and therefore, this observation is more theoretical than practical, it should, however, be mentioned.

8 Conclusion

It was the aim through this research work to articulate specific algorithms, for the well-known and hard to solve problem of FAS, through which one could cover wide spectrum of problem instances—while at the same time, accomplishing aforementioned in such a way where the reader would not be daunted by large number of procedures with at least one question: Which one to choose?

Algorithms were presented (GreedyFAS, BergerShorFAS, MonteCarloFAS, and ExactFAS) through important details, in a concise manner, with original or improved pseudocode, concluding with valuable information and references for further study.

Presented procedures have been critically discussed and comparatively interpreted. This will enable choosing of an appropriate algorithm through a "glance," since clutter is minimized, and the reader is not overwhelmed with information.

This paper will have at least a twofold benefit. For a researcher/practitioner, it will give him a head start by offering a few good algorithms and referring him to additional resources as per ones need. And for a learner, it will introduce him to the problem of FAS and adjacent topics in an easy enough manner, with more complex material at his fingertips.

Notes and Comments. The author of the paper would like to make a special mention of Linux Mint distribution[12] and TeXstudio[13] which were used during preparation of this conference paper—let hard work not be left unnoticed.

References

1. Baharev A, Schichl H, Neumaier A, Achterberg T (2021) An exact method for the minimum feedback arc set problem. ACM J Exp Algorithm 26:1–28. https://www.mat.univie.ac.at/~neum/ms/minimum_feedback_arc_set.pdf
2. Berger B, Shor PW (1990) Approximation algorithms for the maximum acyclic subgraph problem. In: SODA '90: Proceedings of the first annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, pp 236–243. https://dl.acm.org/doi/10.5555/320176.320203
3. Bessy S Bougeret M Krithika R Sahu A Saurabh S Thiebaut J Zehavi M (2021) Packing arc-disjoint cycles in tournaments. Algorithmica 83:1393–1420
4. Cormen TH, Leiserson CE, Rivest RL, Stein C (2009) Introduction to algorithms. MIT Press, third edn
5. Eades P, Lin X, Smyth W (1993) A fast and effective heuristic for the feedback arc set problem. Inf Process Lett 47(6):319–323
6. Festa P, Pardalos PM, Resende MGC (1999) Feedback set problems. In: Handbook of combinatorial optimization. Springer US, pp. 209–258
7. Gabow HN, Galil Z, Spencer T, Tarjan RE (1986) Efficient algorithms for finding minimum spanning trees in undirected and directed graphs. Combinatorica 6(2):109–122
8. Garey MR, Johnson DS (1979) Computers and Intractability: a guide to the theory of NP-completeness. W. H. Freeman & Co. https://dl.acm.org/doi/book/10.5555/578533

[12] Web page of the distribution available at: https://linuxmint.com/.

[13] Editor web page available at: https://www.texstudio.org/.

9. Hecht M (2017) Exact Localisations of Feedback Sets. Theory Comput Syst 62(5):1048–1084 May
10. Kahn AB (1962) Topological sorting of large networks. Communications of the ACM 5(11): 558–562
11. Kann V (1992) On the approximability of NP-complete optimization problems. phdthesis, Royal Institute of Technology, Stockholm, Sweden . https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.66.9127&rep=rep1&type=pdf
12. Karp RM (1972) Reducibility among combinatorial problems. In: Complexity of computer computations. Springer US, pp 85–103
13. Kudelić R (2016) Monte-carlo randomized algorithm for minimal feedback arc set problem. Appl Soft Comput 41:235–246
14. Kudelić R (2022) Feedback arc set: a history of the problem and algorithms. Springer
15. Kudelić R, Ivković N (2019) Ant inspired monte carlo algorithm for minimum feedback arc set. Expert Syst Appl 122:108–117
16. Kudelić R, Rabuzin K (2020) Dealing with intractability of information system subsystems development order via control flow graph reducibility. In: Proceedings of the 2020 3rd international conference on electronics and electrical engineering technology. ACM, pp 62–68
17. Lawler E (1964) A comment on minimum feedback arc sets. IEEE Trans Circuit Theory 11(2):296–297
18. Misra P, Raman V, Ramanujan MS, Saurabh S (2013) A polynomial kernel for feedback arc set on bipartite tournaments. Theory Comput Syst 53(4):609–620 Feb
19. Ramachandran V (1988) Finding a minimum feedback arc set in reducible flow graphs. J Algorithms 9(3):299–313
20. Schwikowski B, Speckenmeyer E (2002) On enumerating all minimal solutions of feedback problems. Discrete Appl Math 117(1–3):253–265
21. Simpson M, Srinivasan V, Thomo A (2016) Efficient computation of feedback arc set at web-scale. Proc VLDB Endowment 10(3):133–144

Prototype of a Simulator for Hemorrhage Control During Tactical Medical Care for Combat Wounded

Sonia Cárdenas-Delgado, Chariguamán Quinteros Magali Fernanda, Pilca Imba Wilmer Patricio, and Mauricio Loachamín-Valencia

Abstract Tactical emergency medical services are outpatient care provided in hostile situations by specially trained professionals. The care and attention to the sick wounded or injured patients is of vital importance to save lives. Disasters, combats, and conflicts are high-acuity events that occur in chaotic, hostile, and austere environments. Our country's armed forces need to strengthen the training of military personnel and form a team of combat nurses, or their equivalent, tactical medical operators. The objective of this work was to develop the prototype of a virtual simulator that included a virtual environment, several 3D objects, and a virtual task. The developed prototype will allow military personnel to be trained to control hemorrhages during tactical assistance to a combat wounded. For interaction and immersion, VR devices were used. The results obtained showed that the virtual task and the devices used in the developed prototype could be a useful and complementary tool for permanent training. The participants reported that they had fun and learned a lot by doing the virtual task because they were completely immersed and focused.

Keywords Virtual reality · Simulator · Hemorrhages · Military training

S. Cárdenas-Delgado (✉) · C. Q. M. Fernanda · P. I. W. Patricio · M. Loachamín-Valencia
Universidad de Las Fuerzas Armadas ESPE, Av. General Rumiñahui S/n, Sangolquí 171-5-231B, Ecuador
e-mail: secardenas@espe.edu.ec

C. Q. M. Fernanda
e-mail: mfchariguamanq@ejercito.mil.ec

P. I. W. Patricio
e-mail: wppilcai@ejercito.mil.ec

M. Loachamín-Valencia
e-mail: mrloachamin@espe.edu.ec

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_6

1 Introduction

Pre-hospital care is an important aspect when providing a medical emergency service. Care and attention are vitally important to the ill, wounded, or injured patient. Disasters, combats, and conflicts are low-frequent events of high-acuity that occur in usually chaotic, hostile, and austere environments. Health care personnel must provide immediate and timely care to the wounded or injured patients. Taking care of medical emergencies will safeguard lives and mitigate the damage of the injuries.

Tactical emergency medical services (TEMS) is outpatient care provided in hostile situations by specially trained professionals. Tactical support is applied in special operations teams of military and civilian personnel. TEMS also encompass the provision of preventive, urgent, and emergency medical care during mission-driven, extended duration, and high-risk law enforcement special operations [16].

The armed forces ensure the security of the country; therefore, their personnel may be involved in war situations and conflicts when carrying out their military operations and/or administrative activities in the military units. Military personnel and special operations teams also face new and challenging situations, including perpetrators armed with military-grade weapons, hostage rescues, entrenched subjects, toxic hazards associated with clandestine drug labs, as well as organized and armed opposing forces.

Trained military personnel must be trained to provide comprehensive emergency care, for this they must have knowledge of the tactical environment and develop skills related to patient assessment and medical treatment in hostile and austere conditions.

The armed forces of our country need to strengthen the training of military personnel of different army branches and form a team of combat nurses or their equivalent as tactical medical operators. However, training soldiers are very costly and time-consuming as it involves, such as expenses related to transporting personnel to specialized grounds and facilities, acquiring specialized equipment, spending on supplies, among others.

Virtual reality (VR) has become more than just a concept in video games, books, and movies. The armed forces of the world are implementing the use of VR in many training sectors due to its innumerable benefits. VR in military training can help place soldiers in a potentially deadly virtual environment. Placing soldiers in such scenarios helps them gain a realistic battlefield experience and gives them the training to act accordingly without exposing them to real-world risks. In addition, the use of simulators has demonstrated the successful transfer of knowledge, skills, and abilities from the simulated environment to the real environment [8, 9].

In this context, we present the first developed component of an emergency tactical medical training simulator for the armed forces of our country. The developed prototype will allow military personnel to be trained to provide emergency tactical support in the control of exsanguinating hemorrhage. The objective is to contribute to the development of skills and competencies of military personnel so that they can control external bleeding during tactical assistance to a combat wounded. The proposal includes a standard model to train and incorporate TEMS in tactical operations through virtual simulation-based training (SBT) [13].

Our hypothesis is that by implementing the first component of the virtual simulator to train military personnel in tactical medical emergencies, it will be possible to develop technical skills to treat exsanguinating hemorrhages and save lives in a harsh and austere environment [4].

In this study, we trained soldiers from different branches of the army through a virtual task under an experimental condition. It consisted of packing a wound and placing a tactical tourniquet on the injured lower limb to stop the bleeding, as well as viewing and interacting with avatars in a hostile virtual environment. In addition, they had to recognize each of the 3D objects in the environment and the emergency material. For visualization, we used an HMD that allows you to feel a more immersive, playful, and realistic experience in the movements you make. For interaction and movement in the virtual environment, the touch control of the HMD was used. After using the simulator, the participants rated 3D sensations and satisfaction [17].

The document is structured as follows. Section 2 contains background of related studies, technology used, and types of simulators. Section 3, describes the development, methods and materials, virtual task, development methodology, hardware and software used. Section 4 describes the study and includes participants, measurements, and a brief description of the procedure performed. Section 5 describes the analysis of results. Section 6 details the conclusions of the experimental study. Finally, Sect. 7 provides an overview of the future works.

2 Background

Applying tactical medicine techniques can be effective in a hostile combat environment to decrease combat personnel fatalities and improve combat effectiveness. According to B.J. Eastridge [5], deaths were generally (90%) caused by blood loss, however, up to 25% of deaths were salvageable. This is vital from the onset of trauma until approximately 10 min after injury [14]. However, most soldiers are not medical professionals and do not have extensive knowledge of human anatomy, which makes it difficult to accurately assess injuries or the severity of bleeding, and therefore, effective treatment is not implemented [15].

Currently, training in first aid and/or tactical medicine is conducted primarily in traditional modes, including training dummies, mock wounds, and animal experiments. Therefore, it is difficult to acquire detailed human anatomical knowledge to perform self-rescue, mutual rescue in a short period of time, and/or control bleeding [7].

Virtual reality (VR) is a technology that allows the generation and visualization of 3D objects and simulated environments to interact with virtual functions. This technology is emerging as a new simulation method. Simulation is becoming a key technique for clinical training in the training of medical professionals. Its applications support different fields.

In recent years, virtual reality (VR) has shown its potential in the medical field as it allows health personnel to be trained in new skills in a safe environment, simulate critical situations, practice different tactical medicine techniques, develop surgical precision skills to save the lives of patients. In addition, this technology allow training new students of medicine and nurses to learn anatomy, practice organization and patient care in different specialties, teaching-learning infection control, internal and external bleeding, other related [11, 12].

Military training in different areas can be just as dangerous as combat. It is said that more soldiers die during training than in combat operations. The use of virtual reality to create a virtual combat environment that saves lives and thus improve negative statistics by improving safety during training. The use of virtual reality in military training for tactical medicine knowledge can produce a new generation of soldiers, saving resources, improving training to face hostile conditions and combat situations [7].

The reviewed literature describes a series of works that develop different types of simulators in which virtual reality is applied. In [6], VR is applied to treat acute and chronic pain in adults and children. The devices were installed in the homes of the patients because they had difficulty traveling to the hospitals due to the distance from their homes or because of the limitations they had. The work [10] describes simulation software for first aid training. The simulator used 3D wound models, PDF material, and instructional videos for training. According to the results obtained, they concluded that the software was of great help to improve skills and abilities in first aid and reinforce the learning of human anatomy.

Moreover, some studies where VR is applied to develop applications related to visual assessment [1–3], and spatial memory these tools were presented as low-cost alternatives for different visual tests. The results demonstrate the validity of their proposals.

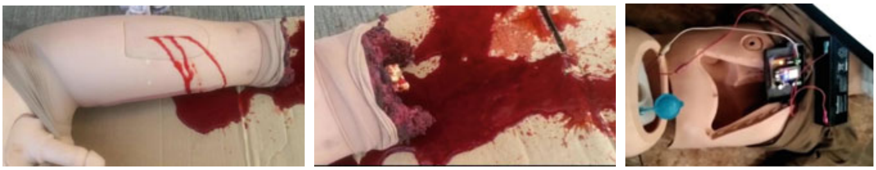
Virtual reality can be a profitable investment in the long term. VR in military training can help place soldiers in a potentially deadly virtual environment. Placing soldiers in such scenarios helps them gain a realistic battlefield experience and gives them the training to act accordingly without exposing them to real-world risks.

In this work, we developed the prototype of a virtual simulator to control hemorrhages that included an interactive and immersive environment using VR devices.

3 Development

The development of the simulator for hemorrhage control (SHC) consisted of two phases: The first phase is a physical-mechanical part, and the second phase is the simulated environment.

The mechanical-physical part was the basis for the development of the simulated virtual part. It is composed of a segment of the lower extremity of the human body. For the development of the simulated environment, a virtual environment was created with five virtual scenarios.



(a) Leg wound (b) Leg with blood flow (c) Mockup with mechanical device

Fig. 1 First Phase: physical-mechanical

3.1 Methods and Materials

Different materials and methods were used to build the mechanical-physical part of the simulator and the simulated environment. Next, each of the constructed phases is described.

First Phase. This phase consisted of building a physical-mechanical model of a wounded leg. The mockup was composed of a segment of the lower extremity of the human body made of silicone and rubber that has an opening with a characterized wound, see Fig. 1. Conduits were placed inside for the passage and exit of the artificial blood. The simulated blood output was driven by a hydraulic pump, which is powered by a 12 V/10 amp power supply. The artificial blood was made with corn glucose, red vegetable dye, impalpable sugar, and unflavored gelatin. This phase was the basis for the model and development of the simulated environment and its components.

Second Phase. This phase consisted of creating a virtual environment. The virtual environment created includes four virtual scenarios, see Fig. 3. Each scenario includes the components and 3D objects agreed upon with specialists for hemorrhage control. The virtual environment and simulated 3D objects created a virtual hostile combat environment.

3.2 Mechanical Task

The built mechanical model was used and control the bleed. The mechanical task consists of packing a wound and placing a tactical tourniquet on the injured lower limb and control bleeding. The participant must ensure that the bleeding has been controlled.

3.3 Virtual Task

Figure 2 shows the scheme of the virtual task of the simulator for hemorrhage control (SHC). The virtual task consists of packing a wound and placing a tactical tourniquet

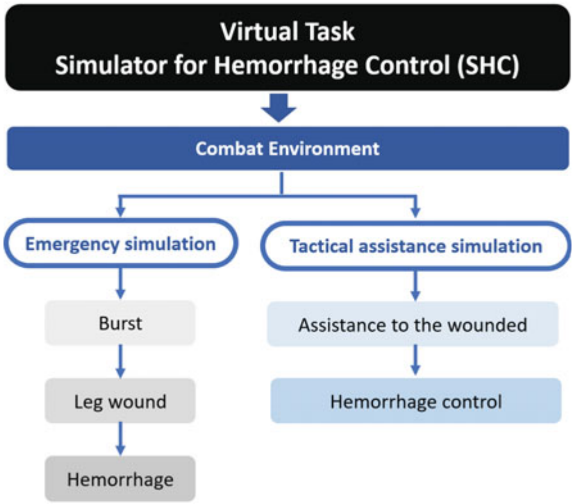


Fig. 2 Scheme of the virtual task

on the injured lower limb and stopping the bleeding, while the participant visualizes and interacts with the avatars in a hostile virtual environment. In addition, they must recognize each of the 3D objects in the environment and the emergency material.

Virtual Environment The virtual environment included five virtual scenarios, 3D objects, and a virtual training task. Furthermore, for the development of the application and integration were used tools and software for modeling, design, and development . 3D Objects were modeled in Blender. The environments configured and programmed with Unity Engine, JavaScript, C#, and the SDKs of each device.

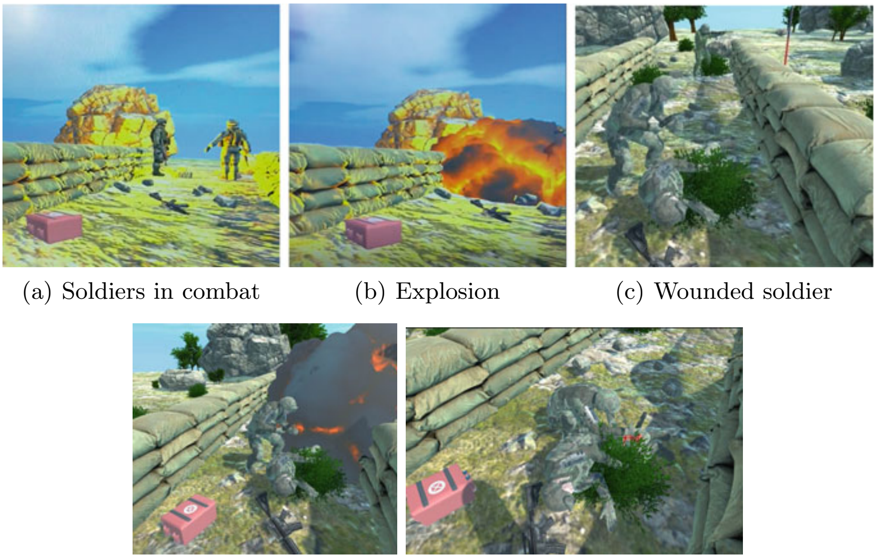
The 3D objects that were included in the environment are combat area, three avatars (injured, combatant, medical assistant), leg, blood, and tourniquet. Each object has been created and textured by the developers.

Figure 3 shows a view of the virtual environments and the objects 3D created and animated to simulate the explosion, soldiers in hostile environment, and assisting soldier with bleeding leg.

The SHC simulator is being developed on a computer with an Intel Core i7 that include a 4GB NVIDIA GeForce GTX-1080 video card. The device used for visualization was an Oculus Rift HMD, which allows a greater sense of immersion within the virtual environment. The virtual scenes were developed in Unity with C# and JavaScript. Blender was used to model the 3D objects.

4 Description of the Study

The study consisted of two experiments. The group of participants in the experimental tests was conformed of military personnel who participated in the II Seminar for updating knowledge in Tactical Medicine. The same group performed the hemorrhage control using both mechanical and virtual simulators.



(a) Soldiers in combat (b) Explosion (c) Wounded soldier

(d) Checking injured (e) Hemorrhage control

Fig. 3 Virtual environment

4.1 Participants

Coordination was made so that the participants of the seminar to update knowledge in Tactical Medicine participate in the experimental study. The total of the group was 52, including 50 men and 2 women. The age of the participants ranged from 26 to 45 years.

4.2 Measurements

Based on the presence questionnaire [18], questions were applied related to the depth of the objects, the interaction with the virtual environment using the devices, and the participants' perception of the immersive experience. Also, questions about the satisfaction perceived by the user in each task performed were used.

4.3 Procedure

The experimental study was developed following a protocol, it can be seen in Fig. 4. Before each session, all the participants were informed about the objectives and

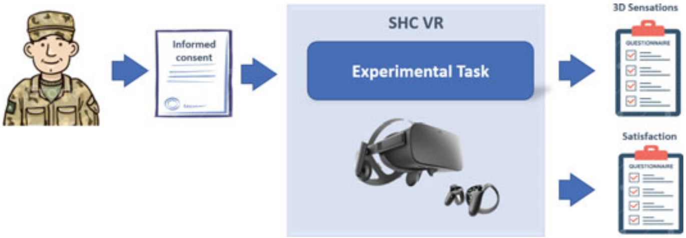


Fig. 4 Protocol

procedures of the study. Also, they signed an informed consent form. They were completely free to withdraw from the study at any time, and the study was conducted in accordance with the principles stated in the Declaration of Helsinki.

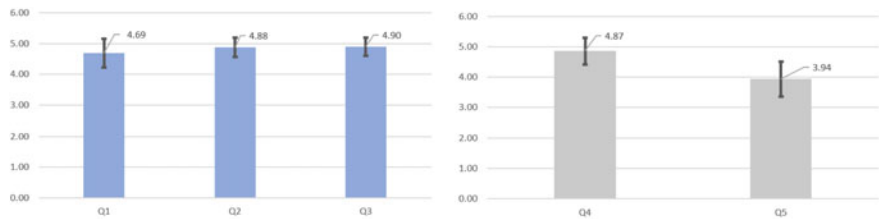
Participants performed two experimental tasks. Each participant was instructed on how to use mechanical task tools (turnstile and scale model) and the devices of the virtual task (HMD and touch controls). Before the task, each participant filled out a form with their personal data. Then, the participants performed the task with the mechanical device. Next, they filled out a satisfaction questionnaire. After that, the participants performed the virtual task and completed the questionnaire about 3D sensations and satisfaction.

5 Results

The data from the study was analyzed using the R-Studio. Analysis of preliminary results is presented in this section. The normality of the data was verified, and based on these results, the pertinent statistical tests were carried out. The Likert scale was applied to each question of the questionnaire.

To determine the results of the 3D sensations, the participants answered questions Q1-Q3 after performing the virtual task. The depth perception score was 4.69/5, the realism of the objects within the virtual environment was 4.88/5, and the sensation of immersion was 4.90/5 (See Fig. 5a). Medians and standard deviation from questions were Q1 (4.69 ± 0.47), Q2 (4.88 ± 0.32), and Q3 (4.90 ± 0.30).The results show that the environment and the components of each virtual scenario have depth, realism, and that the HMD used allowed the participants to feel a completely immersive experience.

Regarding the satisfaction questionnaire, the results of the virtual task reached a score of 4.87/5, and the score of the mechanical task was 3.94/5 (See Fig. 5b). The medians and standard deviation of the questions about satisfaction were Q4 (4.87 ± 0.44) and Q5 (3.94 ± 0.57). These differences obtained show that the virtual task and the devices used in the prototype developed for training in hemorrhage control could be a useful and complementary tool for permanent training. Additionally, 49 of 52 participants reported that they would perform the task again using the virtual



(a) Results of the 3D Sensations (b) Results of the Satisfaction

Fig. 5 Questionnaires

simulator. They also reported that they had fun and learned a lot doing the virtual task because they were completely immersed and focused.

6 Conclusions

The prototype of a simulator for hemorrhage control (SHC) was developed. The SHC consisted of two phases: The first phase is a physical-mechanical part, and the second phase included a virtual environment with five scenarios. For the mechanical task, a mechanical model was built to control bleeding. For the virtual task, visualization and interaction devices, a virtual environment, and the created 3D objects were used.

Regarding the results obtained in the performance comparison between the mechanical and the virtual task. 49 of 52 participants reported that the virtual task was more fun and that they learned more that with the physical-mechanical. Also, they reported that the virtual task allowed greater concentration at the time of training since they were completely immersed.

Regarding the results on 3D sensations and satisfaction, the statistically significant differences found in the two questions were in favor of the virtual task. The participants were satisfied with the proposal shown. They had better depth perception and saw 3D objects within the virtual environment. This study has shown that our SHC simulator prototype could be a useful and complementary tool in the process of training and updating knowledge in tactical medicine for military personnel from the different branches of the army.

The results obtained confirm our hypothesis, since by implementing the first component of the virtual simulator to train military personnel in tactical medical emergencies, it is possible to develop technical skills to control hemorrhages and save lives in a simulated hostile combat environment.

7 Future Works

Our purpose is to continue developing the components of the tactical simulator for emergency medical training for the armed forces of our country. The prototype will allow training military personnel to provide emergency medical care during special operations and hostile situations. In addition, we will continue to develop other components and 3D objects to complete the virtual task and scenarios for hemorrhage control.

Acknowledgements The authors would like to thanks to *Universidad de las Fuerzas Armadas ESPE* for the support provided.

References

1. Cárdenas-Delgado S, Loachamín-Valencia M, Guanoluisa-Atiaga P, Monar-Mejía X (2021) A vr-system to assess stereopsis with visual stimulation: a pilot study of system configuration. In: Artificial intelligence, computer and software engineering advances. pp 328–342
2. Cárdenas-Delgado S, Loachamín-Valencia M, Rodríguez-Reyes B (2022) Vr-test viki: Vr test with visual and kinesthetic stimulation for assessment color vision deficiencies in adults. In: Developments and advances in defense and security. Springer, Singapore, pp 295–305
3. Cárdenas-Delgado S, Loachamín-Valencia M, Rosero-Casa G, Yánez-Lucero F (2022) Design of a low-cost technological alternative tool to assess the amplitude of the visual field, based on an empirical study. In: 2022 17th Iberian conference on information systems and technologies (CISTI). pp 1–6
4. Donley ER, Loyd JW (2018) Hemorrhage control
5. Eastridge BJ, Mabry RL, Seguin P, Cantrell J, Tops T, Uribe P, Mallett O, Zubko T, Oetjen-Gerdes L, Rasmussen TE et al (2012) Death on the battlefield (2001–2011): implications for the future of combat casualty care. J Trauma Acute Care Surgery 73(6):S431–S437
6. Garrett B, Taverner T, McDade P et al (2017) Virtual reality as an adjunct home therapy in chronic pain management: an exploratory study. JMIR Med Inf 5(2):e7271
7. Gjeraa K, Møller T, Østergaard D (2014) Efficacy of simulation-based trauma team training of non-technical skills. a systematic review. Acta Anaesthesiologica Scandinavica 58(7):775–787
8. Glassberg E, Nadler R, Erlich T, Klien Y, Kreiss Y, Kluger Y (2014) A decade of advances in military trauma care. Scandinavian J Surgery 103(2):126–131
9. Holcomb JB, McMullin NR, Pearse L, Caruso J, Wade CE, Oetjen-Gerdes L, Champion HR, Lawnick M, Farr W, Rodriguez S et al (2007) Causes of death in us special operations forces in the global war on terrorism: 2001–2004. Annals of Surgery 245(6):986
10. Hu X, Liu L, Xu Z, Yang J, Guo H, Zhu L, Lamers WH, Wu Y (2022) Creation and application of war trauma treatment simulation software for first aid on the battlefield based on undeformed high-resolution sectional anatomical image (Chinese visible human dataset). BMC Med Educ 22(1):1–10
11. Javaid M, Haleem A (2018) Additive manufacturing applications in medical cases: a literature based review. Alexandria J Med 54(4):411–422
12. Javaid M, Haleem A (2020) Virtual reality applications toward medical field. Clin Epidemiol Global Health 8(2):600–605
13. Lele A (2013) Virtual reality and its military utility. J Ambient Intell Humanized Comput 4(1):17–26
14. Peng Y, Lyu L, Ma B (2020) Advances in the research of application of virtual reality technology in war trauma treatment training. Zhonghua Shao Shang za zhi= Zhonghua Shaoshang Zazhi= Chin J Burns 36(6):515–518

15. Qin H, Liu D, Chen S, Lyv M, Yang L, Bao Q, Zong Z (2020) First-aid training for combatants without systematic medical education experience on the battlefield: establishment and evaluation of the curriculum in China. Mil Med 185(9–10):e1822–e1828
16. Rinnert KJ, Hall WL (2002) Tactical emergency medical support. Emerg Med Clin 20(4):929–952
17. Sharma JP, Salhotra R (2012) Tourniquets in orthopedic surgery. Indian J Orthop 46(4):377–383
18. Witmer BG, Singer MJ (1998) Measuring presence in virtual environments: a presence questionnaire. Presence 7(3):225–240

Automating Systematic Literature Reviews with Natural Language Processing and Text Mining: A Systematic Literature Review

Girish Sundaram and Daniel Berleant

Abstract *Objectives*: An SLR is presented focusing on text mining-based automation of SLR creation. The present review identifies the objectives of the automation studies and the aspects of those steps that were automated. In so doing, the various ML techniques used challenges, limitations, and scope of further research are explained. *Methods*: Accessible published literature studies primarily focus on automation of study selection, study quality assessment, data extraction, and data synthesis portions of SLR. Twenty-nine studies were analyzed. *Results*: This review identifies the objectives of the automation studies, steps within the study selection, study quality assessment, data extraction, and data synthesis portions that were automated, and the various ML techniques used challenges, limitations, and scope of further research. *Discussion*: We describe uses of NLP/TM techniques to support increased automation of systematic literature reviews. This area has attracted increase attention in the last decade due to significant gaps in the applicability of TM to automate steps in the SLR process. There are significant gaps in the application of TM and related automation techniques in the areas of data extraction, monitoring, quality assessment, and data synthesis. There is, thus, a need for continued progress in this area, and this is expected to ultimately significantly facilitate the construction of systematic literature reviews.

Keywords Systematic literature review · Text mining · Automation

G. Sundaram (✉) · D. Berleant
University of Arkansas at Little Rock, Little Rock, AR 72204, USA
e-mail: gsundaram@ualr.edu

D. Berleant
e-mail: jdberleant@ualr.edu

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_7

1 Background

In this section, we describe the motivation of our work beginning with a brief overview of systematic literature reviews (SLRs) especially about the study selection, study quality assessment, data extraction, and data synthesis phases within SLRs. We then also review the existing prior arts to get more insights on the gaps that exist in this field.

A systematic review is one of the numerous types of reviews [1] and is defined as [2] "a review of the evidence on a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant primary research and to extract and analyze data from the studies that are included in the review." The methods used must be reproducible and transparent.

Figure 1 illustrates that systematic reviews are considered to provide the highest quality of evidence in the area of evidence based medicine (EBM). While SLRs are the norm in the field of EBM and healthcare, Kitchenham and Charters [4] provided the framework and guidelines on how SLRs can be used in other fields like software engineering.

Preparing an SLR can be both time consuming and expensive [5, 6]. The time problem is further accentuated by the fact that SLRs become outdated, making their timely completion a quality factor. Shojania et al. [7] show that the median lifetime of an existing review until it needs updating is 5.5 years. It is apparent that the current SLR process needs augmentation to speed up the process of creating them.

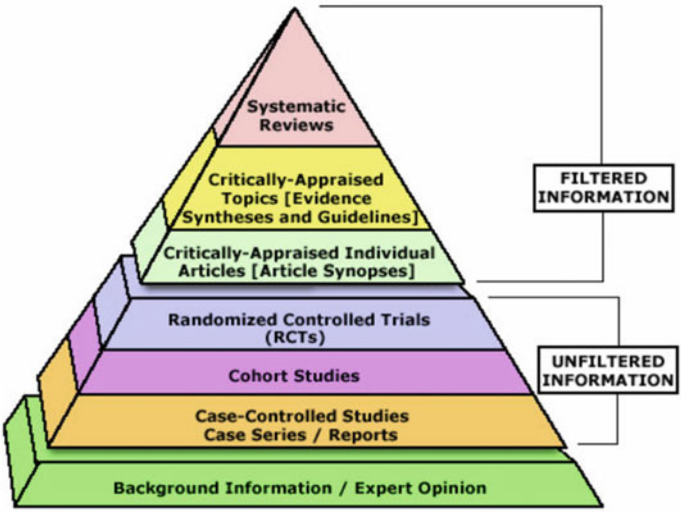


Fig. 1 Systematic reviews (based on Glover et al. [3])

Table 1 Key steps in systematic literature reviews (based on Kitchenham and Charters [4])

ID	Category	Step	Synonyms
SLR1	Defining a review	Commissioning a review	
SLR2		Defining the research questions	
SLR3		Determining a protocol for the review	
SLR4		Evaluating the protocol for the review	
SLR5	Conducting the review	Identification of research	Literature search, search string development
SLR6		Selection of studies	Citation screening
SLR7		Assessing study quality	Selection review
SLR8		Data extraction and monitoring	
SLR9		Data synthesis	
SLR10	Reporting the review	Specifying dissemination mechanisms	
SLR11		Formatting the main report	
SLR12		Evaluating the report	

Specific phases of SLR development such as identification of relevant studies, data collection, extraction, and synthesis have been found to require time consuming and error prone manual effort [8].

NLP and text mining have been used increasingly in the recent past to analyze and automate steps in the SLR process. This paper performs an SLR on the current state of the art. One objective of performing this SLR is to identify specific steps where there has been considerable activity and where there is a scope for further research. We have adapted Table 1 from Kitchenham and Charters [4] to name the steps within the SLR process.

Our primary steps of interest as part of this SLR are SLR5—SLR9. Studies have shown that steps SLR6—SLR9 are often among the most time consuming [9–11].

2 Summary of the Previous Reviews

Here, we describe other reviews of SLR automation to help place the present study in context. Table 2 lists them briefly, followed by additional details.

Jonnalagadda et al. [12] focus on automatic data extraction of critical data elements from full text medical texts as part of the SLR process. They identified 52 potential data elements used in systematic reviews which the authors obtained from standard medical databases/tools such as Cochrane handbook for systematic reviews [21],

Table 2 Related studies

No.	Title	Objective	Reference
1	Automating data extraction in systematic reviews: a systematic review	A systematic review focusing on automatic data extraction prior arts	Jonnalagadda et al. [12]
2	Using text mining for study identification in systematic reviews: a systematic review of current approaches	A systematic review focusing on identifying relevant articles using the title and abstract for reference	O'Mara-Eves et al. [13]
3	Text mining techniques and tools for systematic literature reviews: a systematic literature review	This study presents an SLR in an attempt to understand the application of different TM techniques in facilitating the SLR process. We are interested in identifying the main challenges in the SLRs that can be addressed by applying TM techniques This study presents an SLR in an attempt to understand the application of different TM techniques in facilitating the SLR process. We are interested in identifying the main challenges in the SLRs that can be addressed by applying TM techniques Explains how text mining techniques can contribute to SLR development, focusing on the following text mining categories, namely information extraction, information retrieval, information visualization, classification, clustering, and summarization	Feng et al. [14]
4	Systematic review automation technologies	A systematic review to study the feasibility of automating various phases in an SLR	Tsafnat et al. [15]
5	Toward systematic review automation: a practical guide to using machine learning tools in research synthesis	A guide that can be used by SLR researchers to apply machine learning methods to reduce the overall turnaround time	Marshall et al. [16]

(continued)

Table 2 (continued)

No.	Title	Objective	Reference
6	Moving toward the automation of the systematic review process: a summary of discussions at the second meeting of the International Collaboration for the Automation of Systematic Reviews	Documents various ongoing short-term projects that are carrying out research in the automation of SLR	O'Connor et al. [17]
7	Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews	Documents the outcomes from the conference to improve the overall efficiency of conducting a SLR	Beller et al. [18]
8	Usage of automation tools in systematic reviews	Documents the potential issues that reviewers face when trying to use SLR automation techniques	Van Altena et al. [19]
9	A critical analysis of studies that address the use of text mining for citation screening in systematic reviews	Reviews text mining in the context of systematic literature reviews. More specifically, the focus is on one task within the systematic literature review process. That task is screening citations to determine which ones to include in the review	Olorisade et al. [20]

the consolidated standards of reporting trials (CONSORT) statement, the standards for reporting of diagnostic accuracy (STARD) initiative, PICO [22], PECODR [23], and PIBOSO [24]. The authors concluded that there is no unified data extraction framework that is focused on SLRs, and the prior arts were limited in their scope of the number of data elements (1–7) that were considered in this step. NLP has been limited in its application in this field, and there is considerable scope in further improving its involvement in the data extraction phase of SLRs.

O'Mara-Eves et al. [13] focus on the screening phase of SLR development which is time consuming, and this is further accentuated by the rapid growth in the number of publications in the medical domain. Reviewers have to manually scan through a long list of mostly irrelevant articles that a search yields to identify relevant publications. The paper proposes a solution to semi-automate the screening phase of the SLR process. There are two processes in text mining that can help in screening. One is providing a prioritized list of items, with the ones on the top being the most relevant, that can be used for manual screening by a reviewer. The other is to use machine learning techniques where the system learns from a list of manual classifications of studies as included, or not, and then is able to automatically apply those classifications. They found that both the approaches resulted in reduction of the workload but it was not conclusively proven which method was superior. Some of the research also points out that the performance of the machine learning-based system for relevant article prediction is similar to human efficiency. There is significant potential that this phase can be further improved to reduce the workload in the process [25].

Feng et al. [14] conducted an SLR to identify and classify text mining (TM) techniques to support the SLR process. They classified the various text mining techniques into 6 different categories: information extraction (IE), information retrieval (IR), information visualization (IVi), document classification, clustering, and document summarization. As per their search methodology, out of the shortlisted papers, a majority were focusing on identifying the relevant articles for the study selection stage. They found that the four main applications of TM techniques are (1) visual text mining (VTM), (2) federated search, (3) automated document/text classification, and (4) document summarization. The researchers also attempted to answer the important question of which SLR activities could potentially benefit from TM. There is limited application of TM in the pre-review mapping study as part of the planning phase that Kitchenham and Charters [4] recommend. This is an important phase since the quality of an SLR is directly related to the protocol definition and scoping. There is scope for improvement in the query string development process which is the primary means to locate relevant studies from a variety of sources. Shortlisting and creating the finalized list of articles of interest are the phase where VTM techniques that combine clustering and information visualization have been used by many researchers.

Tsafnat et al. [15] surveyed the literature focusing on automating all aspects of SLRs and found that some of the tasks are fully automatable while many are not. They broke down the SLR process into 4 distinct steps, namely preparation, retrieval, appraisal, and synthesis and write-up and examined the current level and future prospects of automation for these steps. Current research such as global evidence maps and scoping studies can guide in identifying gaps in the work done, helping to provide decision support for reviewers to fine tune and prioritize the research questions [26–28]. They found that despite available tools such as the Cochrane database of systematic reviews and others, creating specific search filters to find relevant items is still a time consuming manual task. There an opportunity to create specialized systems that can understand the nuances of SLR questions and translate them into search filters for efficient identification of relevant prior work. Doing this currently requires specialized expertise in medicine, library science, and standards. Templates such as the ones provided by Cochrane review manager [29] can be a good starting point, and there is ongoing research in this area.

Computational reasoning tasks have not been used extensively in SLRs [30–32]. The associated language bias problem helps define the scope of further research in the application of OCR and NLP to query definition. The application of automatic query expansion (AQE) (synonym expansion, word sense disambiguation, auto-correction, etc.) is part of the search phase of the SLR process. There is little existing work that automates and replicates sequential searching, removing duplicates, or auto-screening of results, which experts use to progressively tune the search parameters based on relevancy of the search results. Related areas include automating snowballing (pursuing references of references) [33] and auto-extraction of important information such as trial features, methods, and outcomes from the texts of shortlisted literature. Overall, the authors conclude that there are significant potential benefits of automating the SLR process using AI/ML.

Marshall et al. [16] review with practical examples how automation technologies can be used, situations where they might help, strengths/weaknesses, and how an SLR team can put these technologies into practice. Although significant work is being done in this area, concerns about accuracy of the current processes limit adoption and highlight the advantage of "human in the loop" automation rather than full automation. Search automation to expedite identification of relevant articles is the most advanced and commercial tools such as Abstrackr, RobotAnalyst, and EPPI reviewer which have been used for secondary screening.

O'Connor et al. [17] and Beller et al. [18] discuss the proceedings of the Int. Collaboration for Automation of Systematic Reviews (ICASR). The authors observed that the number of datasets and tools for automation is increasing, while at the same time, integrating the various tools into a workflow remains a challenge. Most of the tools for screening consider only the abstract and title for classification, but using the full text is complicated by the fact that most of the articles are in PDF format. Acceptance of automation tools is limited due to concerns about accuracy and validity.

Van Altena et al. [19] write about the issues with adoption of automation tools in the SLR process. Candidate tools for their survey conducted were chosen from the "systematic review toolbox" [34] Website, which is a comprehensive list of tools compiled by researchers in the field. The survey results point out that researchers are willing to consider automation and generally feel that they can help. At the same time, there are deterrent factors like poor usability, steep learning curves, lack of support, and difficulty in integrating into a workflow. Another important observation was most of the tools did not have the ability to explain how the results were produced.

Olorisade et al. [20] critically analyzed the various text mining techniques used to augment the SLR process, specifically focusing on the citation screening phase. Certain models like support vector machines (SVMs), Naïve Bayes (NB), and committee of classifiers ensembles were the most commonly used. Current automation research for SLR development focuses on study identification, citation screening, and data extraction using tools such as SLuRp, StArt, SLR-Tool, and SESRA [35–38]. They found missing or insufficient information about details needed for replicating research results such as number of support vectors being used in an SVM model or the number of neurons or hidden layers used. Progress seems slow given the amount of research being done in this area.

There is a need for SLRs focusing on automation of study selection, study quality assessment, data extraction, and data synthesis using TM techniques including NLP, which TM techniques work the best, and performance and accuracy comparisons of human vs AI/TM-driven approaches. Thus, there is a need for an SLR which focused on these questions.

Table 3 Research questions (adapted from van Dinter et al. [40])

No.	Research questions (RQ)
RQ1	Which phase of the SLR process is the focus of automation?
RQ2	Which TM/AI techniques (Table A2 in supplementary material) have been employed for automation?
RQ2A	Which TM/AI models/algorithms have been explored for automation?
RQ2B	Which TM/AI models/algorithms were the most heavily explored?
RQ2C	What evaluation methodology and metrics are used?
RQ3	How will the adoption of TM techniques facilitate SLRs?
RQ4	What is the improvement from employing TM techniques over a manual process?
RQ5	What are the open challenges and solution directions?

3 Identifying the Articles for this SLR

As we have seen in the related work section, there is a need for SLRs that focus on SLR automation using TM and AI. The field of AI/TM is growing rapidly, and such an SLR is expected to help accelerate further studies in this domain. We follow the guidelines explained by Kitchenham and Charters [4] which specifically highlight the need for a well-defined protocol to reduce bias, increase rigor, and improve reproducibility.

3.1 Research Questions

From the software engineering perspective, we aim to find all relevant information for the SLR process. The protocol for review defined by Gurbuz and Tekinerdogan [39] has been adapted here for this SLR.

Table 3 documents the research questions to address. These questions are relevant from a text mining/ML perspective when building a model to extract meaningful insights from a text corpus.

3.2 Search Scope

The SLR scope needs to include the time frame for publication and the sources from where the articles are sourced. Based on the literature that we have seen so far a reasonable time frame was found to be 20 years so we decided to adhere to this. The year 2000 was kept as beginning of the search time frame, and 2021 was fixed as the cutoff date. The language considered for inclusion was English, and the reference type was journals, conference, workshops, symposiums, and book chapters since so many existing SLRs restrict the reference type to these kinds of articles.

3.3 Search Method

We used automated search to retrieve relevant articles from publication repositories. Google Scholar was used as the primary retrieval tool, and we also used snowballing in our search process to identify related relevant articles [41].

3.4 Search String

We tried various combinations of keywords and Boolean operators to construct a search string. This was optimized iteratively to retrieve the maximum number of relevant results by manually conducting a number of trial searches. There is a character limitation of 256 chars in Google Scholar which required tuning the search string accordingly. This led to the following search query as the basis of the search:

("systematic literature review" OR "systematic review") AND ("Automation") AND ("Data Mining" OR "Text Mining" OR "NLP").

3.5 Criteria for Selection

Based on the research questions, we have defined for this SLR, and we formulated the inclusion and exclusion criteria to fine tune the results from the search process. We followed a two phase screening process. Studies that satisfied the following inclusion criteria were included in the first phase:

- Publication year is after 2000
- The study describes a TM/AI/NLP process to support SLR automation (either complete automation or specific phases of SLR)
- If the SLR was a review of other SLRs, then the reviewed articles are evaluated separately
- If multiple versions are available, the most recent one is used

The output from the 1st stage screening process was then manually reviewed (2nd stage) with input from an external consultant to ensure that the results were relevant. Any record marked as doubtful, meaning that it is not clear if it is a relevant article, was discussed for a final decision regarding its inclusion.

The following exclusion criteria were applied to further fine tune the result set.

- The language of publication was not English
- Full text of the article was not accessible
- No empirical results were presented
- Not focused on automation of SLRs specifically
- Discusses TM/NLP/AI techniques but not in the context of SLR automation
- Focused on a particular commercial or open source tool (e.g., Abstrackr)

3.6 Quality Assessment

We assessed the quality of the studies in this stage that were being considered for inclusion. The papers in the search list were read in full text, and a standard assessment criteria were applied to ascertain the quality score. We adapted the quality assessment criteria used by Feng et al. [14] which in turn was developed from the checklist provided by Dybå et al. [42] and Nguyen-Duc et al. [43] See Table A1 in the supplementary material. Only articles scoring sufficiently high in quality, as detailed below, were included.

3.7 Data Extraction

We used a data extraction template to collect all necessary information from the selected primary list of literature to facilitate an in-depth analysis based on our research questions specified in Table 3. The main extraction elements used in that template are listed below. The detailed data extraction template is provided in Table A3 (supplementary material). We adapted the classification of TM methods as specified by Feng et al. [14] for use in the form for data extraction.

- Title
- Passed inclusion criteria
- Year of publication
- Authors
- SLR steps automated
- Level of automation
- Type of review
- TM methods used (category)
- TM model/algorithm information
- TM model evaluation methodology used (if specified)
- Evaluation metrics used
- TM methods used as additional reviewer
- Deep learning or AI used?
- Sampling techniques used

- Overall results/conclusions (stated by authors)
- Performance gain over manual methods provided?

3.8 Data Analysis

As part of the data analysis stage, we analyzed the data extracted from the relevant articles from the earlier step, answering the research questions mentioned in Table 3. For RQ1, we identified the specific phase of an SLR that is the focus of automation in in the article. For RQ2, we categorized the TM/AI techniques used by the researchers and documented them appropriately using the TM categories in Table A2 (supplementary material). In addition, we also documented which techniques were found to be most appropriate in the study and the evaluation metrics used in the process. For RQ3 and RQ4, we gathered information on the adoption of TM techniques for SLRs and potential performance gain observed over traditional methods. Finally for RQ5, we collected information on the open challenges that were observed either directly in an SLR or topics that seemed inadequately addressed in the SLR. Scores were assigned based on quality assessment criteria (Table A1 in supplementary material) as follows:

1. Quality assessment criteria were completely addressed (score was 3).
2. Quality assessment criteria were addressed with moderate gaps (score was 2).
3. Quality assessment criteria were addressed with considerable gaps (score was 1).
4. Quality assessment criteria were not addressed (score was 0).

The list of the papers was finalized after filtering on the quality assessment score as explained below.

3.9 Data Synthesis

Data synthesis is the process of interpreting extracted data to answer the research questions of an SLR. Since each paper might have different naming standards for defining the objectives, naming the algorithms used, models used, etc., we synthesized the data collected using synonyms to gain information on the data patterns.

Using the methods described in the previous sections, this section collates and summarizes the results.

Search Results and Identification of Studies. Using the search strategy mentioned in the above sections, we first searched the six electronic databases for relevant studies using queries as described in Sect. 3.4. The process to identify and finalize the set of studies is depicted in Fig. 2. Table 4 shows the digital libraries that were searched and the final number of papers that were shortlisted from each library after applying the finalization process shown in Fig. 2.

For Google Scholar, the initial set returned during the search using the criteria mentioned in Sect. 3.4 was 15,500. Google Scholar would not provide records past 1000. We observed that the relevancy of the search dataset diminished rapidly after the first 600 records and decided not to screen the remaining records based on this observation. The search results were then screened for relevancy (1st and 2nd stage screenings). Full text documents were retrieved only for the screening final set, for detailed quality analysis and to determine the final list of articles to review.

The articles in the screening final set from Table 4 were then subjected to a quality review (QR), explained below to give the QR final set row. We then used the QR final set for a manual snowball (MS) search to ensure that more relevant results, if any, were added to the cumulative final list.

Quality Assessment Results. The QR final set shown in Table 4 was created after subjecting the 2nd stage screening results to a quality assessment review introduced in Sect. 3.6. Only articles having a quality assessment score ≥ 2 were included in the final list. Full texts for all of the articles were analyzed during this process. We followed the same procedure for the manual snowballing exercise as well. The final list is the combination of the QR final set and finalized MS search rows in Table 4. This multi-step process of quality review and screening was designed to result in a final list of articles that were of high quality and the most relevant articles for our research questions.

4 Discussion and Conclusion

In this section, we analyze the results and answer the research questions shown in Table 3. Table A4 in the supplementary material describes these studies [48–73] for reference.

RQ1: Which phase of the SLR process is the focus of automation?

As part of our analysis, we identified 29 studies which were relevant based on the research questions and contribution to automation of the SLR steps mentioned in Table 1. Some studies were focused on automating multiple stages. During our analysis, we derived the following insights.

- 24 studies focused on automating stage SLR6, selection of studies.
- 8 studies focused on automating stage SLR8, data extraction, and monitoring.
- 1 study focused on stage SLR9, data synthesis.
- 1 study focused on stage SLR5, identification of research.

RQ2: Which categories of TM/AI techniques (Table A2, supplementary material) have been employed for automation?

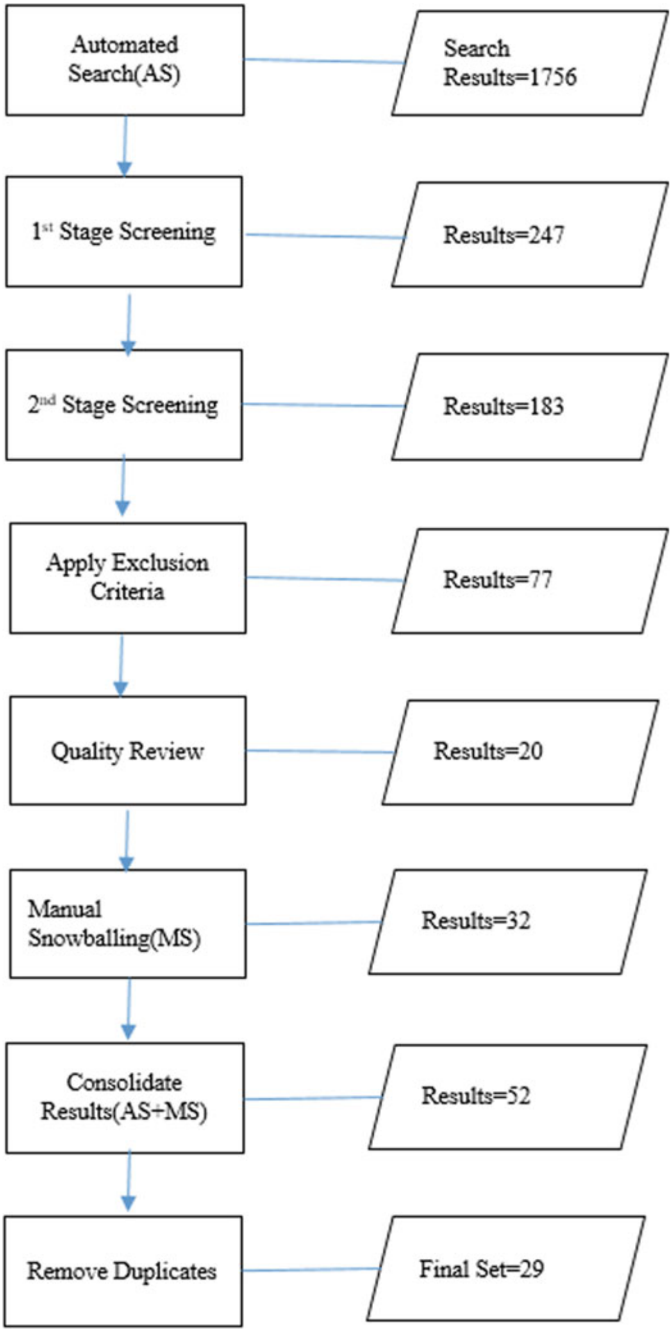


Fig. 2 Selecting and finalizing studies

Table 4 Search results for each digital library

Digital library →	Google scholar	PubMed	Web of science core collection	ACM	IEEE	Science direct	Total
Initial set returned	15,500	143	440	468	60	645	17,256
1st stage screening	63	19	63	77	13	12	247
2nd stage screening	53	19	60	32	11	8	183
Screening final set	28	3	27	14	1	4	77
QR final set	11	2	2	5	0	0	20
Initial MS search	28	3	27	14	0	0	72
Screened MS search	11	2	14	5	0	0	32
Finalized MS search	11	2	14	5	0	0	32
Cumulative list	22	4	16	10	0	0	52
Final (no duplicates) list	10	2	12	5	0	0	29

Twenty-four studies (77%) were related to classification (categorization), two (6%) were related to clustering, two to information extraction (IE), two to information retrieval (IR), and one (3%) to summarization. Clearly, the preponderance of the studies used TM/AI methods for classification.

RQ2A: Which TM/AI models/algorithms have been explored for automation?

We analyzed the data extracted from the finalized set of studies which is documented in Table A4 (supplementary material).

RQ2B: Which TM/AI models/algorithms were the most heavily explored?

As indicated in Fig. 3, SVM, logistic regression, Naïve Bayes, and random forests are the most frequently used algorithms.

RQ2C: Which evaluation methodologies and metrics have been used?

An evaluation metric as "a metric quantifies the performance of a predictive model [44]. This typically involves training a model on a dataset, using the model to make predictions on a holdout dataset not used during training, then comparing the predictions to the expected values in the holdout dataset." We found that cross validation was the most frequently used evaluation methodology.

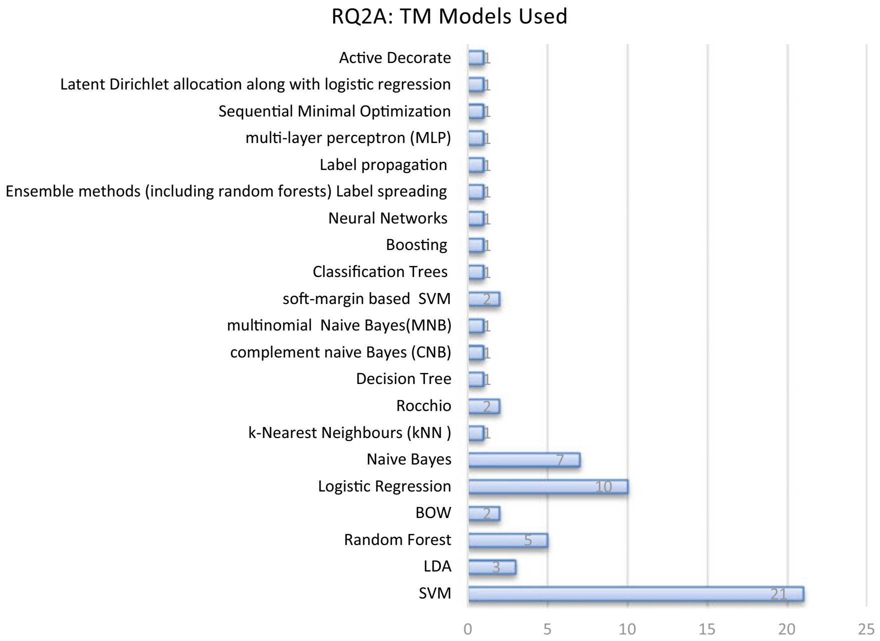


Fig. 3 Most frequently used TM models for experimentations

RQ3: How will the adoption of TM techniques facilitate SLRs?

As Marshall et al. [16] mention in their paper, the most frequently used application of TM/NLP techniques in the SLR field is text classification and data extraction. We arrived at the same conclusion as part of our SLR as well. Classification methods are generally used to classify the paper in question as relevant or not based on the research questions, the SLR is attempting to address. Data extraction on the other hand is used to extract important portions from the SLR to get data for a particular variable or attribute of interest. For example, extracting the PICO elements from an SLR is a very common application of data extraction. TM techniques appear to have the potential to significantly impact the quality and speed of SLR development.

RQ4: What is the improvement from employing TM techniques over manual processing?

There was mention of improvement due to using the NLP/TM model. Wallace et al. [45] mention that they were able to reduce the number of citations to be screened by 40–50% without excluding any relevant ones. Pham et al. [46] achieved workload reduction in the range of 55–63% with the number of missed studies in the range of 0–1.5%. Norman et al. [47] found that the main meta-analysis for each systematic review can be reliably performed with an estimation error of 1.3% average after screening around 30% of the candidate articles.

RQ5: What are the open challenges and solution directions?

As mentioned in explaining RQ1 above, the majority of the research has been focused on SLR6 (selection of studies), and a distant second is SLR8 (data extraction and monitoring). Important SLR activities such as SLR5 (identification of research, including search string development and literature search), SLR7 (assessing study quality), and SLR9 (data synthesis) have been less scrutinized. We also found that there is need and opportunity for continued development of artificial intelligence techniques in the SLR process.

4.1 Summary of Conclusions

In this SLR, we have described the use of NLP/TM techniques in the area of automation of SLR development. This area has been active for the last decade and continues to attract more research. As mentioned in RQ5, there are significant gaps in the application of TM and other automation techniques in the areas of data extraction, monitoring, quality assessment, and data synthesis. AI in the SLR automation process has experienced a recent surge of exploration, and there is a need to continue this due to the promise of improved automation and all the benefits flowing therefrom.

4.2 Future and Ongoing Research

The primary objective of conducting this systematic literature review was to find the current state of the art in the field regarding applying NLP, TM, and AI techniques to automate specific steps in an SLR. After conducting this SLR, we found that there are significant opportunities to use NLP to assist SLR and more specifically in the data extraction phase. We are currently conducting research on using various NLP techniques to assist in extraction of PICO [22] data elements from randomized control trial free text articles and summarizing the overall clinical evidence.

Acknowledgements Publication of this work was supported by the National Science Foundation under Award No. OIA-1946391. The content reflects the views of the authors and not necessarily the NSF. The authors are grateful to Deepak Sagaram, MD, for consulting on the list of articles regarding their relevance for inclusion and exclusion.

Supplementary Material

The supplementary material, including data Tables A1–A4, may be obtained at https://dberleant.github.io/papers/sundber-supp.pdf.

References

1. Systematic reviews. Georgetown University Medical Center. https://guides.dml.georgetown.edu/systematicreviews
2. Systematic reviews (2001) CRD's guidance for those carrying out or commissioning reviews. CRD Report Number 4 (2nd edn). NHS Centre for Reviews and Dissemination, University of York
3. Glover J, Izzo D, Odato K et al (2006) EBM pyramid and EBM page generator. Trustees of Dartmouth College and Yale University
4. Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report EBSE-2007-01. Keele University. https://docs.edtechhub.org/lib/EDAG684W
5. Allen IE, Olkin I (1999) Estimating time to conduct a meta-analysis from number of citations retrieved. JAMA 282(7):634–635. https://doi.org/10.1001/jama.282.7.634
6. Petticrew M, Roberts H (2006) Systematic reviews in the social sciences: a practical guide. Blackwell Publishing Co., Malden
7. Shojani KG, Sampson M, Ansari MT et al (2007) How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med 147(4):224–233. https://doi.org/10.7326/0003-4819-147-4-200708210-00179
8. Marshall C, Kitchenham B, Brereton P (2018) Tool features to support systematic reviews in software engineering. E-Informatica Softw Eng J 12(1):79–115. https://doi.org/10.5277/e-Inf180104
9. Khangura S, Konnyu K, Cushman R et al (2012) Evidence summaries: the evolution of a rapid review approach. Syst Rev 1:10. https://doi.org/10.1186/2046-4053-1-10
10. Ganann R, Ciliska D, Thomas H (2010) Expediting systematic reviews: methods and implications of rapid reviews. Implementation Sci 5:56. https://doi.org/10.1186/1748-5908-5-56
11. Featherstone RM, Dryden DM, Foisy M et al (2015) Advancing knowledge of rapid reviews: an analysis of results, conclusions and recommendations from published review articles examining rapid reviews. Syst Rev 4:50. https://doi.org/10.1186/s13643-015-0040-4
12. Jonnalagadda SR, Goyal P, Huffman MD (2015) Automating data extraction in systematic reviews: a systematic review. Syst Rev 4:78. https://doi.org/10.1186/s13643-015-0066-7
13. O'Mara-Eves A, Thomas J, McNaught J et al (2015) Using text mining for study identification in systematic reviews: a systematic review of current approaches. Syst Rev 4:5. https://doi.org/10.1186/2046-4053-4-5
14. Feng L, Chiam Y, Lo SK (2017) Text-mining techniques and tools for systematic literature reviews: a systematic literature review. In: 24th Asia-Pacific software engineering conference (APSEC 2017). https://doi.org/10.1109/APSEC.2017.10
15. Tsafnat G, Glasziou P, Choong MK et al (2014) Systematic review automation technologies. Syst Rev 3(74). https://doi.org/10.1186/2046-4053-3-74
16. Marshall IJ, Wallace BC (2019) Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. Syst Rev 8:163. https://doi.org/10.1186/s13643-019-1074-9
17. O'Connor AM, Tsafnat G, Gilbert SB et al (2018) Moving toward the automation of the systematic review process: a summary of discussions at the second meeting of the international collaboration for the automation of systematic reviews (ICASR). Syst Rev 7:3. https://doi.org/10.1186/s13643-017-0667-4
18. Beller E, Clark J, Tsafnat G et al (2018) Making progress with the automation of systematic reviews: principles of the international collaboration for the automation of systematic reviews (ICASR). Syst Rev 7:77. https://doi.org/10.1186/s13643-018-0740-7
19. Van Altena AJ, Spijker R, Olabarriaga SD (2019) Usage of automation tools in systematic reviews. Res Syn Meth 10:72–82. https://doi.org/10.1002/jrsm.1335

20. Olorisade BK, de Quincey E, Brereton OP et al (2016) A critical analysis of studies that address the use of text mining for citation screening in systematic reviews. In: EASE '16: proceedings of the 20th international conference on evaluation and assessment in software engineering. ACM, Limerick, pp 1–11. https://doi.org/10.1145/2915970.2915982
21. Higgins J, Green S (2011) Cochrane handbook for systematic reviews of interventions version 5.1.0. The Cochrane Collaboration. http://community.cochrane.org/handbook.
22. Richardson WS, Wilson MC, Nishikawa J et al (1995) The well-built clinical question: a key to evidence-based decisions. ACP J Club 123(3):A12–A13
23. Dawes M, Pluye P, Shea L et al (2007) The identification of clinically important elements within medical journal abstracts: patient–population–problem, exposure-intervention, comparison, outcome, duration and results (PECODR). Inform Prim Care 15(1):9–16
24. Kim S, Martinez D, Cavedon L et al (2011) Automatic classification of sentences to support evidence based medicine. BMC Bioinform 12(Suppl 2):S5
25. Razavi A, Matwin S, Inkpen D et al (2009) Parameterized contrast in second order soft co-occurrences: a novel text representation technique in text mining and knowledge extraction. In: 2009 IEEE international conference on data mining workshops, pp 71–6
26. Bragge P, Clavisi O, Turner T et al (2011) The global evidence mapping initiative: scoping research in broad topic areas. BMC Med Res Methodol 11(92). https://doi.org/10.1186/1471-2288-11-92
27. Snilstveit B, Vojtkova M, Bhavsar A et al (2016) Evidence and gap maps—a tool for promoting evidence informed policy and strategic research agendas. J Clin Epidemiol 79:120–129. https://doi.org/10.1016/j.jclinepi.2016.05.015
28. Arksey H, O'Malley L (2005) Scoping studies: towards a methodological framework. Int J Soc Res Meth 8:19–32
29. RTC Collaboration. Review Manager (RevMan) 4.2 for Windows. The Cochrane Collaboration, Oxford (2003)
30. Tsafnat G, Coiera E (2009) Computational reasoning across multiple models. J Am Med Info Assoc 16(6):768–774
31. Sim I, Detmer DE (2005) Beyond trial registration: a global trial bank for clinical trial reporting. PLoS Med 2(11):e365
32. Sim I, Tu SW, Carini S et al (2014) The ontology of clinical research (OCRe): an informatics foundation for the science of clinical research. J Biomed Inf 52:78–91. https://doi.org/10.1016/j.jbi.2013.11.002
33. Greenhalgh T, Peacock R (2005) Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ 331(7524):1064–1065. https://doi.org/10.1136/bmj.38636.593461.68
34. Marshal C, Sutton A, O'Keefe H et al (2022) The systematic review toolbox. http://www.systematicreviewtools.com.
35. Bowes D, Hall T, Beecham S (2012) SLuRp : a tool to help large complex systematic literature reviews deliver valid and rigorous results. In: Proceedings of the 2nd international workshop on evidential assessment of software technologies—EAST '12, pp 33–36
36. Hernandes E, Zamboni A, Fabbri S et al (2012) Using GQM and TAM to evaluate StArt—a tool that supports systematic review. CLEI Electr J 15(1):2. http://www.scielo.edu.uy/pdf/cleiej/v15n1/v15n1a03.pdf
37. Fernández-Sáez AM, Bocco MG, Romero FP (2010) SLR-Tool—a tool for performing systematic literature reviews. In: ICSOFT 2010—proceedings of the 5th international conference on software and data technologies, pp 157–166
38. Molléri JS, Benitti FBV (2015) SESRA: a web-based automated tool to support the systematic literature review process. In: EASE '15: proceedings of the 19th international conference on evaluation and assessment in software engineering, pp 1–6. https://doi.org/10.1145/2745802.2745825
39. Gurbuz HG, Tekinerdogan B (2018) Model-based testing for software safety: a systematic mapping study. Software Qual J 26:1327–1372. https://doi.org/10.1007/s11219-017-9386-2

40. Van Dinter R, Tekinerdogan B, Cagatay C (2021) Automation of systematic literature reviews: a systematic literature review. Inf and Software Tech 136:106589. https://doi.org/10.1016/j.infsof.2021.106589
41. Wohlin C (2014) Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: EASE '14: proceedings of the 18th international conference on evaluation and assessment in software engineering.. ACM, pp 1–10. https://doi.org/10.1145/2601248.2601268
42. Dybå T, Dingsøyr T (2008) Empirical studies of agile software development: a systematic review. Inf Softw Tech 50(9):833–859
43. Nguyen-Duc A, Cruzes DS, Conradi R (2015) The impact of global dispersion on coordination, team performance and software quality—a systematic literature review. Inf and Softw Tech 57:277–294
44. Brownlee J, Tour of evaluation metrics for imbalanced classification. https://machinelearningmastery.com/tour-of-evaluation-metrics-for-imbalanced-classification
45. Wallace BC, Trikalinos TA, Lau J et al (2010) Semi-automated screening of biomedical citations for systematic reviews. BMC Bioinformatics 11(1):55. https://doi.org/10.1186/1471-2105-11-55
46. Pham B, Jovanovic J, Bagheri E et al (2021) Text mining to support abstract screening for knowledge syntheses: a semi-automated workflow. Syst Rev 10:156. https://doi.org/10.1186/s13643-021-01700-x
47. Norman CR, Leeflang M, Porcher R et al (2019) Measuring the impact of screening automation on meta-analyses of diagnostic test accuracy. Syst Rev 8:243. https://doi.org/10.1186/s13643-019-1162-x
48. Dickson K (2017) Systematic reviews to inform policy: institutional mechanisms and social interactions to support their production. Dissertation. University College London. http://discovery.ucl.ac.uk/id/eprint/10054092/1/KD_PhD_FinalAugust2018_Redacted.pdf
49. Turing A (1950) Computing machinery and intelligence. Mind LIX (236):433–460. https://doi.org/10.1093/mind/LIX.236.433
50. Mo Y, Kontonatsios G, Ananiadou S (2015) Supporting systematic reviews using LDA-based document representations. Syst Rev 4:172. https://doi.org/10.1186/s13643-015-0117-0
51. Cohen AM, Ambert K, McDonagh M (2012) Studying the potential impact of automated document classification on scheduling a systematic review update. BMC Med Inform Decis Mak 12:33. https://doi.org/10.1186/1472-6947-12-33
52. Callaghan MW, Müller-Hansen F (2020) Statistical stopping criteria for automated screening in systematic reviews. Syst Rev 9:273. https://doi.org/10.1186/s13643-020-01521-4
53. Miwa M, Thomas J, O'Mara-Eves A et al (2014) Reducing systematic review workload through certainty-based screening. J Biomed Inf 51:242–253. https://doi.org/10.1016/j.jbi.2014.06.005
54. Basu T, Kumar S, Kalyan A et al (2016) A novel framework to expedite systematic reviews by automatically building information extraction training corpora. arXiv:1606.06424 [cs.IR] (2016). https://arxiv.org/abs/1606.06424
55. García Adeva JJ, Pikatza Atxa JM, Ubeda CM et al (2014) Automatic text classification to support systematic reviews in medicine. Expert Syst with Appl 41(4):1498–1508. https://doi.org/10.1016/j.eswa.2013.08.047
56. Ros R, Bjarnason E, Runeson P (2017) A machine learning approach for semi-automated search and selection in literature studies. In: EASE '17: proceedings of the 21st international conference on evaluation and assessment in software engineering. Association for Computing Machinery, New York, pp 118–127. https://doi.org/10.1145/3084226.3084243
57. Frunza O, Inkpen D, Matwin S (2010) Building systematic reviews using automatic text classification techniques. In: Proceedings of the 23rd international conference on computational linguistics: poster, vol (COLING '10). Association for Computational Linguistics, pp 303–311
58. Timsina P, Liu J, El-Gayar O (2016) Advanced analytics for the automation of medical systematic reviews. Inf Syst Frontiers 18(2):237–252
59. El-Gayar OF, Liu J, Timsina P (2015) Active learning for the automation of medical systematic review creation. In: 21st Americas conference on information systems (AMCIS). Puerto Rico Aug 13–15. http://aisel.aisnet.org/amcis2015/BizAnalytics/GeneralPresentations/22

60. Halamoda-Kenzaoui B, Rolland E, Piovesan J et al (2021) Toxic effects of nanomaterials for health applications: how automation can support a systematic review of the literature? J of Appl Tox 42(1):41–51. https://doi.org/10.1002/jat.4204
61. Olorisade BK, Brereton P, Andras P (2019) The use of bibliography enriched features for automatic citation screening. J of Biomed Inf 94:103202. https://doi.org/10.1016/j.jbi.2019.103202
62. Bannach-Brown A, Przybyła P, Thomas J et al (2019) Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. Syst Rev 8(1):23. https://doi.org/10.1186/s13643-019-0942-7
63. Bui D, Del Fiol G, Hurdle JF et al (2016) Extractive text summarization system to aid data extraction from full text in systematic review development. J Biomed Inf 64:265–272. https://doi.org/10.1016/j.jbi.2016.10.014
64. Tsafnat G, Glasziou P, Karystianis G et al (2018) Automated screening of research studies for systematic reviews using study characteristics. Syst Rev 7:64. https://doi.org/10.1186/s13643-018-0724-7
65. Norman C (2020) Systematic review automation methods. Université Paris-Saclay, Universiteit van Amsterdam https://tel.archives-ouvertes.fr/tel-03060620/document
66. Norman C, Leeflang M, Zweigenbaum P et al (2018) Automating document discovery in the systematic review process: how to use chaff to extract wheat. In: Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018). European Language Resources Association (ELRA), Miyazaki, Japan. https://aclanthology.org/L18-1582
67. Karystianis G, Thayer K, Wolfe M et al (2017) Evaluation of a rule-based method for epidemiological document classification towards the automation of systematic reviews. J of Biomed Inf 70:27–34. https://doi.org/10.1016/j.jbi.2017.04.004
68. Wallace BC, Kuiper J, Sharma A et al (2016) Extracting PICO sentences from clinical trial reports using supervised distant supervision. J Mach Lear Res 17:132
69. Marshall IJ, Kuiper J, Wallace BC (2015) Automating risk of bias assessment for clinical trials. J Biomed Health Inf 19(4):1406–1412. https://doi.org/10.1109/JBHI.2015.2431314
70. Ma Y (2007) Text classification on imbalanced data: application to systematic reviews automation. Dissertation. University of Ottawa
71. Begert D, Granek J, Irwin B et al (2020) Towards automating systematic reviews on immunization using an advanced natural language processing-based extraction system. Can Commun Dis Rep 46(6):174–179. https://doi.org/10.14745/ccdr.v46i06a04
72. Scells H, Zuccon G, Koopman B (2019) Automatic boolean query refinement for systematic review literature search. In: The World Wide Web Conference (WWW '19). Association for Computing Machinery, New York, pp 1646–1656. https://doi.org/10.1145/3308558.3313544
73. Khabsa M, Elmagarmid A, Ilyas I et al (2016) Learning to identify relevant studies for systematic reviews using random forest and external information. Mach Learn 102:465–482. https://doi.org/10.1007/s10994-015-5535-7

Anomaly Detection in Orthopedic Musculoskeletal Radiographs Using Deep Learning

Nabila Ounasser, Maryem Rhanoui, Mounia Mikram, and Bouchra El Asri

Abstract In this paper, we investigate anomaly detection in orthopedics musculoskeletal radiographs using deep learning. We examine thirteen models from the most powerful neural network families: generative adversarial networks (GANs), autoencoders (AEs) and convolutional neural network (CNN). The main goal is to detect anomalies in musculoskeletal radiographs using the Stanford Musculoskeletal Radiographs (MURA) dataset. The results of the examined models were compared to several recent studies. Generally, CNN models achieve the best score of 0.822 which is a very promising result, competitive for expert radiologists performance.

Keywords Anomaly detection · Orthopedic · Radiographs · Deep learning · Autoencoders · Generative adversarial networks · Convolutional neural network

1 Introduction

Musculoskeletal abnormalities are the most common pathology, abiding suffering and disability, which makes detecting anomalies in radiographs correctly a crucial task in the medical world. Analyzing X-ray for diagnosing orthopedics diseases (bone malformation, tumors, fractures) is time consuming and requires qualified experts. Therefore, developing a computer-aided diagnosis system to detect anomalies has become an attractive domain in X-ray imaging. In this way, artificial intelligence (AI) is explored by researchers, distinctly deep learning, to provide diagnostics assistance to radiologists to improve the quality of patient care and reducing time of diagnosis.

N. Ounasser (✉) · M. Rhanoui · B. El Asri
IMS Team, ADMIR Laboratory, Rabat IT Center, ENSIAS, Mohammed V University, Rabat, Morocco
e-mail: nabilaounasser81@gmail.com

M. Rhanoui · M. Mikram
Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_8

In this context, several studies examine deep learning-based anomaly detection (AD) methods [6, 20] that ease anomalies detection in X-ray images more correctly [1, 4, 12, 26] Deep learning-based anomaly detection includes several approaches, and we can categorize these approaches as (1) supervised; (2) semi-supervised; and (3) unsupervised. Supervised models use labeled datasets to train algorithms to classify data or predict outcomes accurately. Contrary to the unsupervised approach which is characterized by its flexibility, it is adjustable because it does not require labeled data [20]. In this study, we will explore both of the methods to have a rich comparative study of musculoskeletal anomaly detection. Several models have been investigated in this direction such as convolutional neural network (CNN) [5, 18], autoencoders (AEs) and generative adversarial networks (GANs) [8, 20, 25]. In this work, we investigate how we can improve anomaly detection in X-ray images tasks quality. To this end, we demonstrate how unsupervised and supervised methods, such as generative adversarial networks (GANs), autoencoders (AE) and convolutional neural network (CNN), may be drawn on orthopedic anomaly detection on X-ray images. We will explore MURA dataset, the largest public radiographic image datasets. The contribution through this work can be summarized as follows:

- Review the examined deep learning models, approaches and architectures.
- Preprocessing techniques are applied to prepare radiographs for further analysis.
- Combine some models to configure an ensemble model to improve performance.
- Comparison and analysis study are built based on the results of the examined models.

The rest of the paper is designed as follows: Sect. 2 lays out a brief summary of the related work, and Sect. 3 presents our paper's methodology. Section 4 describes materials used, dataset and implemented models and analyzes and compares the results. Finally, Sect. 5 concludes the findings of this paper.

2 Related Works

Analyzing X-ray is a common medical way for diagnosing orthopedics diseases: bone malformation, tumors and fracture. However, anomaly musculoskeletal detection on radiographs is time consuming and requires qualified experts. Therefore, many researchers work on developing a computer-aided diagnosis system to search for anomalies in X-ray imaging. This section presents the previous studies on anomaly detection task. Many researchers have trained convolutional neural networks on bone X-ray images. In [1], authors used transfer learning techniques including fine-tuning or through feature extractor musculoskeletal abnormality detection on X-ray images. Some recent works used deep anomaly detection methods such as generative adversarial networks (GANs) (GANomaly [2, 3], AlphaGAN [22], BiGAN [9], GAAL [16]) and the deep convolutional autoencoder (AE) [7, 17] to improve the anomaly detection task. Researchers tried to improve the performance of those mentioned models by changing their components. GANomaly was introduced by Ackay et al.

[2, 3] in 2019; it is improved by several extensions example skip connections [3] that utilize an AE that converts the entries into a latent space and reconstructs them in this space. Spahr et al. [26] innovate a new semi-supervised anomaly detection approach, called the self-taught anomaly detection model, capable to work without any prior domain knowledge to detect invisible anomalies. Simonyan and Zisserman [25] propose a Res-UnetGAN model based on GAN architecture and practiced to MURA. The network they introduced contains two components: generator and discriminator. The ResNet50 plays the role of the encoder in the generator part [13] that gets the potential feature vector by extracting features of normal samples. Uzunova et al. [28] adopt the AE approach, precisely VAE as a model to process 2D and 3D CT medical images. To analyze the obtained results, they calculated the MSE reconstruction loss as the abnormality score. In Table 1, we summarize some of the recent works on anomaly detection in musculoskeletal radiographs.

Table 1 Summary of recent works of anomaly detection in musculoskeletal radiographs

[Refs.] (Year)	Dataset	Approach	Results
[25] (2021)	MURA dataset	Propose an unsupervised anomaly detection approach Res-UnetGAN based on a GAN architecture made up of ResNet50 and Unet	Res-UnetGAN: 0.92 GANomaly: 0.81 Skip-GANomaly: 0.90 CVAE-GAN based: 0.86 EGBAD: 0.80
[26] (2021)	MURA dataset: upper limb images	Propose self-taught anomaly detection, adopting encoder network trained on the semi-supervised multimodes anomaly detection	Self-taught: 0.78
[8] (2020)	MURA dataset: hands images	Comparative study between GAN and AE models on anomaly detection	CAE: 0.57 VAE: 0.48 DCGAN: 0.53 BiGAN: 0.54 AlphaGAN: 0.60
[11] (2020)	MURA dataset	Propose GnCNNr model which adopts the principle of normalization, including group normalization, weight normalization and cyclic learning rate planner to improve the model performance measures	DenseNet: 0.879 Inceptionv2: 0.888 GnCNNr: 0.899
[18] (2020)	MURA dataset	Introduce MuRAD, a tool developed to automatize the detection of anomalies in X-ray images of bones, based on a convolutional neural network (CNN)	DenseNet161: 0.84 InceptionV3: 0.78 VGG11: 0.83
[19] (2020)	MURA dataset containing images of humerus	Computer-based diagnosis (CBDs) model based on DenseNet201 and InceptionV3 models, their utility is to distinguish between abnormal and normal samples	DenseNet201: 87.15 InceptionV3: 86.11 Ensemble: 88.54

3 Methodology

In this section, we represent the different approaches we had explored in this study. **GAN** Generative adversarial network (GAN), introduced by Goodfellow [10], is an innovative unsupervised deep learning method to generate fake data from noise; this part is guaranteed by the generator. Then, we have the second discriminator component that deals with the classification of the real and generated data (Fig. 1).

From GAN family, we will examine **BiGAN** presented by Donahue et al. [9] as a combination of the GAN architecture family and an encoder, which distinguishes it to learn from the data space x to the latent space z. The difference in BiGAN resides in the discriminator component that must discriminate between the data pair X and the latent variable Z. **AlphaGAN** [22] extends the GAN architecture by adding a third component autoencoder. Unlike BiGAN, AlphaGAN has a direct connection from encoder to generator.

Autoencoders Autoencoders are deep neural networks composed of two symmetrical deep layers used to reproduce the input data at the output layer. The first network takes care of extracting the characteristic features of the input, and the second network reproduces the output based on the extracted features.

For AE, we will use **VAE**. The variational autoencoder is a powerful AE characterized by regularized learning that allows it to avoid overfitting and ensure that the latent space can take possession of good properties for a successful generative process.

CNN A convolutional neural network is an anticipatory neural network. It is characterized by its sequential design which allows it to learn and retain the features of an input and leave the other features. This method optimizes the cost and computation time. This architecture contains many stacked convolutional layers, each capable of recognizing more sophisticated patterns (Figs. 2 and 3).

Various CNNs were examined, **VGG16 and VGG19** proposed by Simonyan and Zisserman [24] show a promising performance, and it is because of its main feature

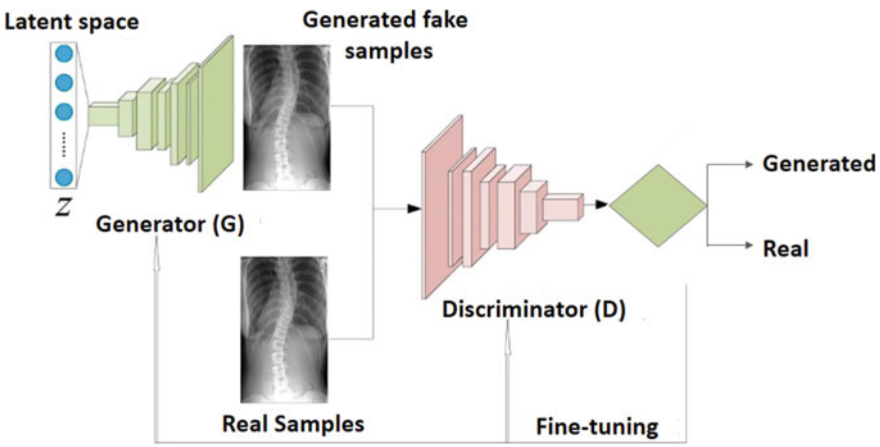


Fig. 1 GAN's architecture

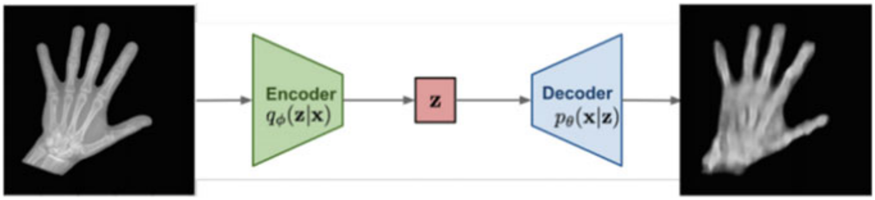


Fig. 2 Autoencoder's architecture [8]

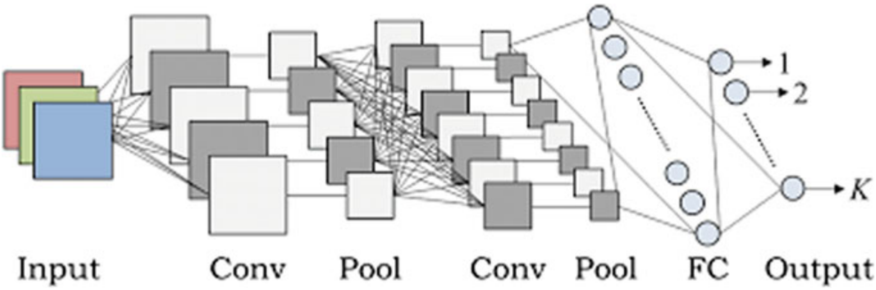


Fig. 3 CNN's architecture [14]

of concatenation of multiple convolutional layers of $k \times k$ filters. This network model has many variations depending on the number of convolutional layers that the model has. In our case, we chose to use the VGG19 and VGG16 variant, i.e., with 19 and 16 deep layers. **DenseNet169** is contributed by Huan et al. [15]. DenseNet is based on the CNN architecture, but the advantage is that we have a deeper model containing a large number of convolutional layers, thus improving the accuracy and the computation efficiency. He et al. contribute **RestNet50** [13]. It has two main advantages. First, it facilitates training through its residual learning. Second, it is distinguished by connecting a number of layers, which allows the layers to easily optimize the underlying residual mapping $H(x)$. In our study, we will use 50-layer variation. Szegedy et al. [27] introduced in 2016 **InceptionV3**; it is considered a true improvement of the Inception family of networks. Its particularity comes from the multi-combination of different filter sizes that make the approach more adaptable to different variations. **MobileNetV2** introduced by Sandler et al. [23] to improve the efficiency of mobile models. What is new in the MobileNetV2 architecture is its inverted residual structure where the input and output layers of the residual block are very thin which makes the computation optimal with promising results. **Ensemble modeling** implemented an ensemble model by combining all three models (VGG16, VGG19, InceptionV3). **Conv2D** is exploring CNN in three approaches. The first is a binary classification normal/abnormal radiograph with a single sigmoid neuron in the output layer. The second approach is to split the two predictions, evaluating the seven body parts and then distinguishing between normal or abnormal radiographs. The final is to create 14 classes, seven body parts * normal or abnormal, and it has 14 softmax neurons $i + n$ the output layer.

4 Experiment

4.1 Dataset

In this study, we will explore the public dataset MURA[21]. MURA is the largest collection of radiographic images with over 40,000 X-ray images, including 9045 normal and 5818 abnormal musculoskeletal radiographic studies of the different organs of the human body, namely the upper limb, including the shoulder, humerus, elbow, forearm, wrist, hand and finger, as shown in Fig. 1.

The dataset is provide by the Stanford Program for Artificial Intelligence in Medicine: https://stanfordmlgroup.github.io/competitions/mura/.

4.2 Preprocessing

First, for convolutional models, we constructed a generator, and we defined the input shape to $224 \times 224 \times 3$. For GAN models, images are resized to 128 pixels on the longer image side while maintaining aspect ratio. Next, we had applied data augmentation. We augmented the images, so we can have more diverse data available which help models to learn high-level features of the dataset that are invariant to normal affine transformations that could be present such as horizontal flips and small rotations of the study that could happen when realizing the radiographic image.

4.3 Experiment Results and Discussion

Looking at the accuracy of the models implemented above, we can see that they performed within the range of 48.3–82%. When comparing our score to the overall score from the Stanford Model and Radiologist's accuracy, we can see that some of the implemented models in this paper accuracy are comparable. There are several factors that affect models performance especially architectural approach, layer design, padding, shape, normalization, activation, loss function, optimizer, batch size, learning rate, pooling and output layer can affect the accuracy between the models. After multiple tuning, trying to have an effective result was our goal. Most of our models had at least ten layers and were generally computationally expensive. Training models for entire days sometimes, running on basic hardware or laptop configurations is probably time prohibitive on MURA dataset. Detailly, Table 2 shows that DenseNet169, ResNet50, and VGG16 perform the best and VAE performs the worst by a large margin. The DenseNet and ResNet50 models achieved the best overall performance in all models. In addition, VGG16 obtained the second best results. Some models examined have not achieved a good performance, and this could be a result of using the feature extractor. This reflection is due to the non-change of the con-

Table 2 Performance comparison of trained models on MURA dataset

Category	Models	AUC
AE	VAE	0.48
GAN	BiGAN	0.52
	AlphaGAN	0.64
CNN	VGG16	0.78
	VGG19	0.70
	InceptionV3	0.73
	Ensemble modeling	0.72
	DenseNet169	0.82
	ResNet50	0.82
	MobileNetV2	0.74
	Conv2D (Binary)	0.72
	Conv2D (7 + 2 Class)	0.72
	Conv2D (14 Class)	0.64

volutional extinction layers during processing. This results in the non-compatibility of the high-level features with the MURA data. Since we are dealing with images, two-dimensional convolution layers (Conv2D) model iterates over each pixel with a kernel. Pooling reduces the number of pixels available to the next layer in the model. While running the MobileNet model, we found average pooling to be more effective than max pooling. We also found that deeper models perform better than shallow and wide models. Looking at the accuracy of our models above, we can see that our models are within the range of 0.80–48%. When comparing our score to the overall score from the Stanford Model and Radiologists' accuracy, we can see that some examined models in this paper: Conv2D, DenseNet, ResNet50, MobileNet2D, InceptionV3, VGG16, VGG19 and ensemble model accuracy are comparable. When comparing MobileNetV2, DenseNet169, VGG16, VGG19, and Conv2D models, we found the MobileNetV2 implementation was the better. Part of this can be attributed to MobileNet's lightweight design to be less computational. As we predicted, the results of our ensemble model (VGG16, VGG19, InceptionV3) are promising, as it has demonstrated a better result than an individual configuration. For GANs, we observe AUC values smaller than 70%, AlphaGAN achieves 60.7% and BiGAN 54.9%. Wherefore, they still some missing pieces in the current approach of using GANs for anomaly detection that could be solved to improve anomaly detection systems. To our expectation, we did not expect to get a higher accuracy score compared to the Stanford overall score. Looking ahead, we think the most significant change we can make for our next work is to find the most suitable and efficient combination between GAN, AE and CNN that could give the most performed model for anomaly detection (Fig. 4).

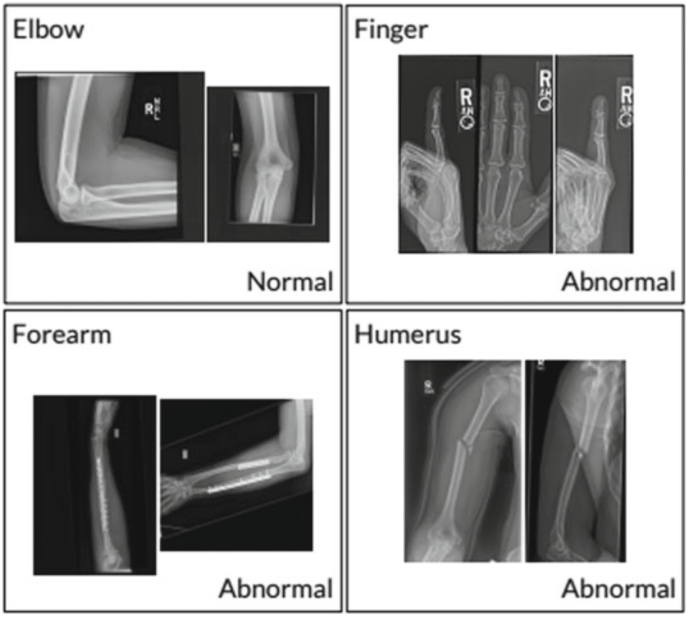


Fig. 4 Dataset description [21]

5 Conclusion

In our experiments, we compare and analyze several models against each other for anomaly detection tasks. We cannot deny our disappointment with the performance of some models, especially GAN, which for the moment cannot be integrated into a computer-aided diagnostic system, but we can take advantage of their usefulness to reduce diagnostic time and minimize errors. Moreover, we can observe the potential of approaches, as well as the possibility of building an ensemble model that could perform better. Therefore as part of the future work, we want to introduce a different training approach by introducing an ensemble model, by combining the three families studied in this work in one architecture, that could learn specific features for detecting musculoskeletal abnormalities.

References

1. Abreu Dias Dd (2019) Musculoskeletal abnormality detection on X-ray using transfer learning
2. Akcay S, Atapour-Abarghouei A, Breckon TP (2018) GANomaly: Semi-supervised anomaly detection via adversarial training. In: Asian conference on computer vision. Springer, pp 622–637

3. Akçay S, Atapour-Abarghouei A, Breckon TP (2019) Skip-GANomaly: skip connected and adversarially trained encoder-decoder anomaly detection. In: 2019 International joint conference on neural networks (IJCNN). IEEE, pp 1–8
4. Alaoui Belghiti K, Mikram M, Rhanoui M, Yousfi S (2023) Deep learning based multi-task approach for neuronal cells classification and segmentation. In: Proceedings of eighth international congress on information and communication technology. Springer
5. Ananda A, Ngan KH, Karabağ C, Ter-Sarkisov A, Alonso E, Reyes-Aldasoro CC (2021) Classification and visualisation of normal and abnormal radiographs; a comparison between eleven convolutional neural network architectures. Sensors 21(16):5381
6. Chalapathy R, Chawla S (2019) Deep learning for anomaly detection: a survey. arXiv:1901.03407
7. Chen Y, Zhang J, Yeo CK (2019) Network anomaly detection using federated deep autoencoding gaussian mixture model. In: International conference on machine learning for networking. Springer, pp 1–14
8. Davletshina D, Melnychuk V, Tran V, Singla H, Berrendorf M, Faerman E, Fromm M, Schubert M (2020) Unsupervised anomaly detection for X-ray images. arXiv:2001.10883 (2020)
9. Donahue J, Krähenbühl P, Darrell T (2016) Adversarial feature learning. arXiv:1605.09782
10. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2020) Generative adversarial networks. Commun ACM 63(11):139–144
11. Goyal M, Malik R, Kumar D, Rathore S, Arora R (2020) Musculoskeletal abnormality detection in medical imaging using GnCNNr (group normalized convolutional neural networks with regularization). SN Comput Sci 1(6):1–12
12. Harnoune A, Rhanoui M, Mikram M, Yousfi S, Elkaimbillah Z, El Asri B (2021) Bert based clinical knowledge extraction for biomedical knowledge graph construction and analysis. Comput Methods Programs Biomed Update 1:100042
13. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
14. Hidaka A, Kurita T (2017) Consecutive dimensionality reduction by canonical correlation analysis for visualization of convolutional neural networks. In: Proceedings of the ISCIE international symposium on stochastic systems theory and its applications, vol. 2017. The ISCIE symposium on stochastic systems theory and its applications, pp 160–167
15. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708
16. Liu Y, Li Z, Zhou C, Jiang Y, Sun J, Wang M, He X (2019) Generative adversarial active learning for unsupervised outlier detection. IEEE Trans Knowl Data Eng 32(8):1517–1528
17. Matsumoto M, Saito N, Ogawa T, Haseyama M (2019) Chronic gastritis detection from gastric X-ray images via deep autoencoding Gaussian mixture models. In: 2019 IEEE 1st global conference on life sciences and technologies (LifeTech). IEEE, pp 231–232
18. Mehr G (2020) Automating abnormality detection in musculoskeletal radiographs through deep learning. arXiv:2010.12030
19. Namit Chawla NK. Musculoskeletal abnormality detection in humerus radiographs using deep learning
20. Ounasser N, Rhanoui M, Mikram M, Asri BE (2022) Generative and autoencoder models for large-scale mutivariate unsupervised anomaly detection. In: Networking, intelligent systems and security. Springer, pp 45–58
21. Rajpurkar P, Irvin J, Bagul A, Ding D, Duan T, Mehta H, Yang B, Zhu K, Laird D, Ball RL, et al (2017) MURA: large dataset for abnormality detection in musculoskeletal radiographs. arXiv:1712.06957
22. Raza K, Singh NK (2021) A tour of unsupervised deep learning for medical image analysis. Curr Med Imaging 17(9):1059–1077
23. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) MobileNetV2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520

24. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556
25. Song S, Yang K, Wang A, Zhang S, Xia M (2021) A MURA detection model based on unsupervised adversarial learning. IEEE Access 9:49920–49928
26. Spahr A, Bozorgtabar B, Thiran JP (2021) Self-taught semi-supervised anomaly detection on upper limb X-rays. In: 2021 IEEE 18th international symposium on biomedical imaging (ISBI). IEEE, pp 1632–1636
27. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826
28. Uzunova H, Schultz S, Handels H, Ehrhardt J (2019) Unsupervised pathology detection in medical images using conditional variational autoencoders. Int j Comput Assisted Radiol Surgery 14(3):451–461

Steering Data Arbitration on Facial-Speech Features for Fusion-Based Emotion Recognition Framework

Vikram Singh and Kuldeep Singh

Abstract Emotion recognition is one of the computationally complex tasks with a spectrum of real-world applications. In this view, recent years' range of potential strategies is designed based on monolithic learning primarily over a single data modality data source. Though, emotive analytics asserted the capacity of inclusion of additional data modalities for the multifaceted emotion recognition task, with an improved recognition rate. With the evidence of a fusion-based learning strategy, the feature set of multimodal data may be harnessed in an adaptive fusion-based emotion recognition framework. We proposed a fusion-based framework using speech and image features of the reference object for an improved emotion recognition strategy. The role of data arbitration to steer the learning and recognition is highlighted and asserted, with an implicit capacity to handle heterogeneity at learning model and data modality, achieved accuracy equivalent to humanize, e.g., 90.32% recognition rate.

Keywords Data argumentation · Deep learning · Human emotion recognition

1 Introduction

The recognition of human emotion is a fundamental task in several real-world application scenarios. A potentially facial feature-based strategy is expected to consider multiple factors, e.g., face detection, face element observation, facial expressions, speech analysis, behavior (gesture/posture) or physiological signals, and many more for completing the recognition task [1, 2]. A data-centric approach to design accurate and effective emotion recognition requires attaining data learning and further

V. Singh (✉) · K. Singh
National Institute of Technology, Kurukshetra, Haryana 136119, India
e-mail: viks@nitkkr.ac.in

K. Singh
e-mail: kuldeep_32013111@nitkkr.ac.in

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_9

implicit capability within model to adapt to the data or features-related changes. In these settings, a multimodal data scenario emerges as clear winner; with implicit challenges of data heterogeneity is the key challenge to achieve an accurate recognition task.

In recent years, emotive analytics evolves as an interesting research area of work, with a blend of psychological and technological efforts [3, 4]. Though, data arbitration is a pertinent task in this, as data or features suggest that modality may be facing limits in obtaining relevant information that would allow them to derive the most value out of such essential features for the recognition task. A strategy with adaptive data arbitration is pivotal to reconcile heterogeneity and overfitting/under fitting challenges and eventually improved the recognition rate.

The adaption on visual data modality and speech data modality is one key direction for the development of an emotion recognition strategy. Both modalities cater the important features located to the human emotions, though detecting faces within a photo or video and sense expressions by analyzing the relationship between feature points on the face is a complex task. The elements of visual data modality steer the overall emotion recognition due to its reductive nature and range of detection values. With visual data or more specifically facial expression, human emotions are recognized mainly as *Happy, Sadness, Anger, Fear, Surprise, Neutral, and Disgust*, as shown in Fig. 1. The speech data features supplement in the accuracy of the recognition model.

A fusion-based learning model is the recent trend for the development of potential tools and models for data-centric human emotion recognition, due to its capacity to cater heterogeneity of data and overcome the inherent overfitting and under fitting

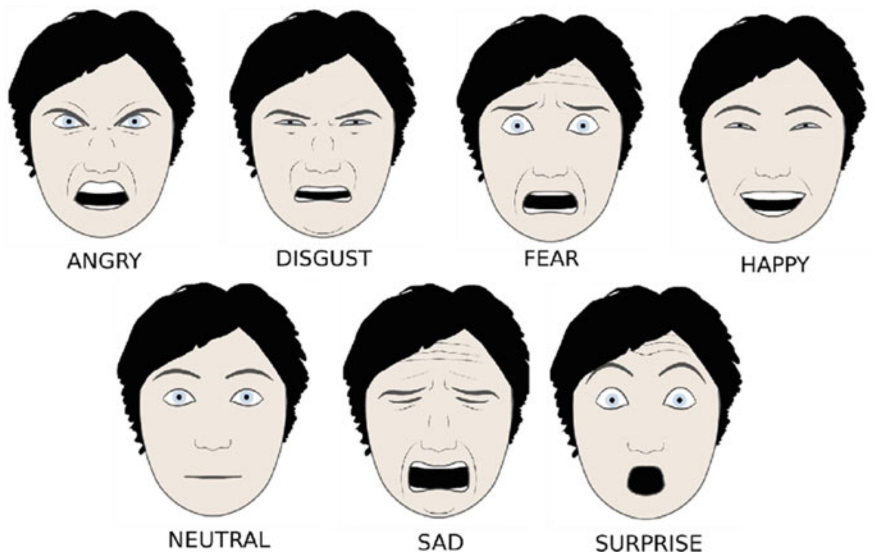


Fig. 1 Facial image object with implicit human emotions

scenario [5–7]. In the view of range of existing fusion-based learning models, selecting an appropriate model is a tedious task.

In this paper, we have adapted lightweight learning models over two modalities, speech and visual, and fused them into a single learning framework. The designed recognition framework drives on lower implicit parameters as adapted learning model for both modalities are customized slightly to align to higher level of recognition accuracy. In this process, identification of potential feature sets, combination of different layers, choices of different datasets, and choices of different base model which will have a direct impact on the combined results are design issues.

We have experimented with the combination of several deep learning-based techniques for both data modalities and analyzed the learning outcomes, i.e., learning rate, recognition rates, accuracy, etc. The customization on the model for each modality is also done to some extent. At the end of design phase, two potential models are fused into a single framework with co-located data arbitration strategies. The data arbitration assists in deriving the relationship between feature points of both data modality, with primary emphasis on facial features and supplemented with the speech feature of same reference objects. Both features set passes through equivalent models to attain learning for emotion labels.

In the paper, we have experimented on the various approaches of facial features-based emotion recognitions, speech-based emotion recognitions, and fusion-based strategy with an accuracy of about 64.3%, 85.6%, and 90.32%, respectively. The recognition rate is similar to human accuracy, i.e., 60–70%. The experimental observations of designed approach over popular datasets, i.e., *FER-2013*, RAVDESS, TESS, and SAVEE for training, validation, and testing our models, further, used *three cross-validation* on all the datasets, as *Train, Test, and Validation.*

1.1 Design Challenges and Research Questions (RQs)

Emotion recognition is a multifaceted computational task, due to intrinsic requirement of learning-based model over cognitive of features set. Even among the emergence of deep learning-based algorithms in recent years, most of the existing research efforts are limited on two design fronts, *(i) they are based on the assumption that increasing the number of layers will also increase the recognition rate, that is hypothetical to an extent* and (ii) *they are unable to deal with multimodal data scenario.* Conceptually, first argument is based on the basis of increased depth of model and affects the recognition rate. The soft clarification for the second reason could be lesser number of experimentations for the facial emotion recognition scenario.

In the views of above, design issues and overall motivation of the proposed work are to overcome the limitation on a fusion-based model and deliver a highly accurate recognition model using facial and speech features set. The following research questions (RQs) are formalized to conduct the design and overall analysis:

RQ I: Which of the deep learning techniques may be kept in a suite to design a novel fusion-based human emotion recognition model?

RQ II: How much stacking must be given to facial features or speech features, in order to get improved recognition rate?

RQ III: How data arbitration plays a pivotal role for facial-based learning models data scenario?

The key motivation of the factor of consideration this work is fascination about how machines or programs recognize human emotions, '*How by using a set of camera or datasets machine learn to recognize corresponding to happy, sad, and other emotions as well'*. In this work, a new custom model for emotion recognition is developed, enabled with existing deep learning models, in the multimodal setting.

1.2 Contribution and Outline

The key contribution is a novel human emotion recognition framework. A fusion-based recognition framework is designed to detect human emotions in the presence over multimodal data, e.g., visual/facial and speech. The framework is designed to stack the learning layers in complex computational settings, mainly due to modeling of features-level relationship from both modalities. The effective fusions of heterogeneous models are framed with a detailed performance analysis. In this work, a new stacking arrangement is contributed after *MobileNetV2* as base model. Dataset formation: In our work, new datasets formation which involves a modified *FER-2013* for facial emotion recognition is adapted. Various argumentation parameters are used to achieve the following task. For speech, a custom model and raw speech samples are used for emotion recognition, and at last a concatenation layer is used for merging the two models input and finally predicting the output.

The paper is organized as Sect. 1 introduced outlines the fundamental aspects of fusion-based emotion recognition with implicit design issue. Section 2 presents the current research work of relevant area. Section 3 presents the proposed approach, i.e., how an appropriate fusion affects the recognition accuracy. Section 4 elaborates details of overall performance evaluation framework and outcome. Section 5 concludes the work presented in the paper and listed the future scopes.

2 Related Work

Convolution neural network (CNN) has shown a great potential in image processing when first arrived in late 1990. Akhdand [1, 2] used eight pre-trained deep convolution neural network (DCNN) models and used transfer learning to avoid training from starch, i.e., freezing all the layers except the last block and used ten cross-validation and famous datasets like *KDRF* and *JAFFE* used for model evaluation. In [3], author adopted the use of *VGG* with regularization and used *SGD* as optimizer and other optimization methods and used the saliency map for visualization. In [4], instead of using deep dense network it uses deep sparse network and inception layers and used seven publicly available datasets. In [5], author uses switching from *Adam* to *SGD (SWATS*); here Adam is adaptive moment estimation and SGD as stochastic gradient descent; i.e., during middle of training, it can take advantage of fast convergence of *Adam* at the beginning, but later we can use *SGD* so that model generalizes well [6].

In this author used *AFF-Wild* 2 dataset and train *CNN* in this model and then tested it on *FER-2013* [7]. Amil et al. used *transfer learning, data argumentation, class weighting, auxiliary data* and also used ensemble with soft voting to archive accuracy of 75.8%. In [8], authors added a *local normalization process* between *CNN* layers that detect *smile* and recognize facial expressions.

Gupta and Vishwamitra [9] combine spatial and temporal features of same reference object located within a video, as features aggregation reduces overfitting problem in *CNN* models, whereas Sinha and Aneesh [9] proposed a similar architecture as *VGG* with doubled *conv* layers and different data argumentation techniques and further extended by Shervin, Mehdi, and Amirali work. A novel attention-based technique has been introduced, leveraging an end-to-end *Convolutional Neural Network (CNN)*, to emphasize crucial facial attributes by employing a *localization network*. This innovative approach yields enhanced recognition results, pushing the boundaries of facial recognition technology [10, 11].

The main processes of an emotion recognition strategy are extraction, selection, and classification. To extract features, Wang [12] employed a *DAE* with five hidden layers, while Khalil [13] examined separate methods for three processes to attain deep learning. Stuhlsatz et al. [14] compared the performance of a GerDA based on deep neural networks (DNNs) with support vector machines (SVM) on identifying speech utterances using emotional dimensions such as arousal and valence [15]. The static acoustic properties were retrieved using a particular preprocessing approach and provide input data to the classifiers, and their findings demonstrated that the *DNN* surpassed the *SVM* in recognizing emotional traits in spoken utterances.

Caridakis et al. [16] employ recurrent neural networks (RNNs) which combine audio formats like MPEG-4 FAPs and data linked to our pitch and its rhythm in order to recognize natural emotional states in terms of activation and valence. This form of neural network implies that previous inputs influence future input processing, giving a framework for dynamic modeling of multimodal data.

Ranganathan et al. [1] use four deep belief networks to extract robust multimodal features for emotion classification in an unsupervised manner, as well as

CDBN models that learn important characteristics of emotions. These models are verified on the *emoFBVP* database, which contains (facial, body gesture, voice, and physiological signal information) and provides improved identification accuracy.

3 Proposed Fusion-Based Emotion Recognition Framework

Fusion-based emotion recognition harnesses the feature set from diverse data modality and based on transfer learning strategy. We propose a multiple stacked layers placed within learning model to achieve this transfer learning for the motion recognition on seven emotion values.

Network Architecture for Modified MobileNetV2 The proposed architecture for the emotional classifier involves number of layers, as each computational layer in the model adds a depth to the delivery of overall computational task. If these layers are increased, the overall capacity of model will be increased, though unnecessarily increasing layers will cause model to over fit and putting too much neurons in single layer. These settings further enhance the overall cost of a model instead of separating it out which is computationally efficient.

Figure 2 illustrates the proposed MobileNetV2-based network for the multilevel emotion recognizer. The proposed strategy has input shape of $224 \times 224 \times 3$ that is resized all the dataset images to $224 \times 224 \times 3$ and added over proposed seven layers at last.

In the proposed architecture, a Dense layer is used to acquire all the information from previous layer then performed activation so that overall values squash between *(0,x)* where *x* is value greater than 0, which is *ReLu* functionality. Interesting point to be note here is that from all the 3 models *MobileNetV2* used *ReLU6* as underlying architecture but we used *ReLu* as common *Activation* in our model.

The proposed model not changed the underlying architecture to match the proposed, because it will change the overall architecture of the predefined model and we have not adapted the *ReLu6* because other model used the underlying architecture as *ReLu* itself. Since layers can cause the overfitting issue so we used the random dropouts of 10, 20 and 50% to reduce the overfitting and used *SoftMax* as classification layer.

Figure 3 illustrates emotion recognition architecture based on the speech feature, with custom 11 learning layers with an input shape of (268, 1). At the first layer, *Conv1D* layer with the 128 filters, 5×5, is kernel or filter size and padding is set to same, to ensure that if outputs have smaller size than input, then simply add padding to it and make the size same as input. The main aim is to extract features from the speech data.

The second layer is of activation; in proposed work, we have used activation as *ReLU* which squashes the values between *(0, x)*. The third layer is a *dropout* layer which was set to 0.2, this means for the random drop of neurons with probability of

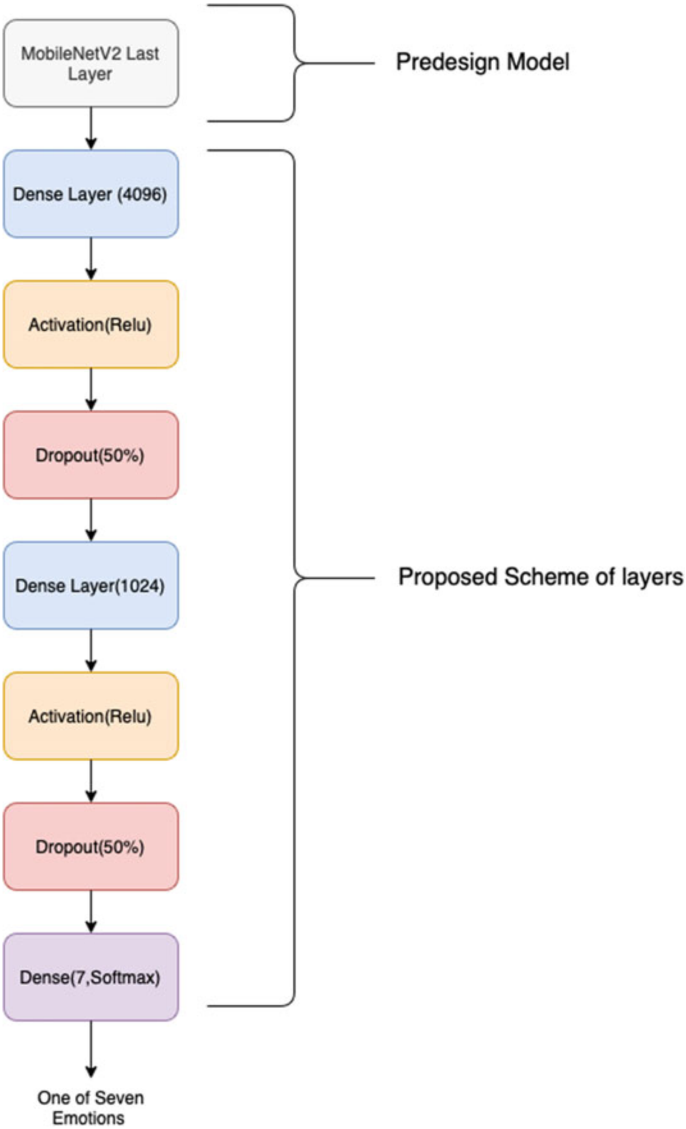


Fig. 2 Proposed MobileNetV2 layer schema

Fig. 3 Custom model for *speech feature-based* emotion recognition

20%. The fourth layer is *Maxpool 1D*, here pool of size is 8, which simply takes max out of eight values and rejects all other values.

The fifth layer consists of *Conv1D* with same 128 filters and five kernels. The Sixth layer is the second activation in model which is also *ReLU*. The seventh layer is *MaxPool1D* with a pool size 5 which is same as the fifth layer. The eighth layer consists of *Conv1D* but has 64 filters and 5 will be the kernel size. The ninth layer is also activation layer of type *ReLU*. The next layer is a dropout layer which has a value of 0.1 which simply means probability of 10% which simply drops 10% of neurons randomly, and at last flatten layer will be used to reduce the multi-dimensional data into single dimension.

All features which are used for speech model are exactly same as features used in speech section, i.e., *MFCC, MEL, Contrast, Tonnetz, and Chroma.* But here we need to reshape the audio sample to twice its size.

In fusion-based system, as shown in Fig. 4, a *concatenation layer* is used to harness the input of same shape, in our case for the architectures mentioned above; we have added two new layers as output layers and both layers are dense in nature with number of neurons which will be seven in both layers. Then after that we have combined all four parameters, i.e., two model structures defined above and two output parameters. During *compilation stage,* we have used *Adam* as well as *RmsProp* as a *loss* function below is the combined structure for the fusion stage.

4 Experimental Analysis

4.1 Data Settings

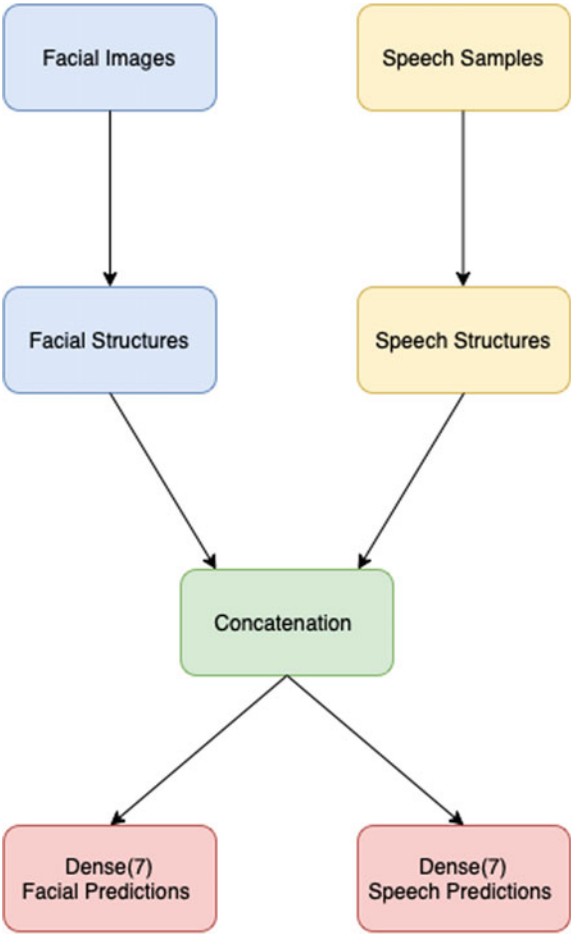
All studies were carried out on a computer equipped with an *Intel i5-8300H 2.5* GHz processor and 16 GB of RAM, with no additional devices such as GPU, *TensorFlow,* and *OpenCV* for pertained models and dataset preprocessing.

The primary source of data is taken from *facial emotion recognition* (*FER-2013*) dataset with located images crawled over Google Search in an uncontrolled environment. The dominant features on these images are set on over seven human emotions with each image (size of 48×48). The key challenge, here, is inherent data imbalance, e.g., dataset is quite imbalance for '*Happy*' emotion having lots of images as compared to '*Disgust*' emotion. Figure 5 (*left* part) illustrated the overall imbalance on the data objects, for each training and test image set.

4.2 Data Arbitration for Emotion Recognition Accuracy

A traditional data argumentation is a data preparation strategy either focuses either on balancing data objects within dataset or enhancing the data objects adding the

Fig. 4 Proposed architecture for fusion-based emotion recognition



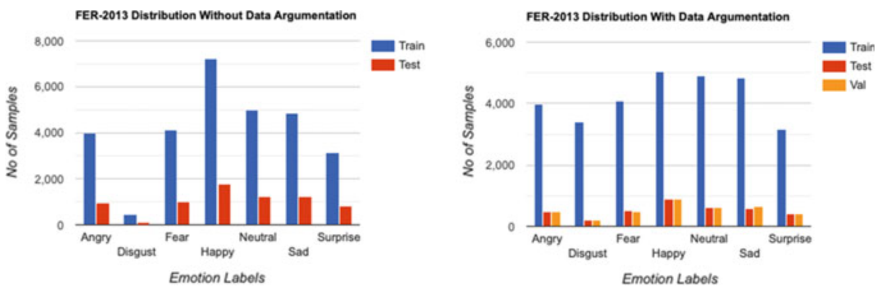


Fig. 5 Data instances imbalance (*left*) and data with data argumentations (*right*)

additional data objects using via synthetic data records. Data arbitration is an implicit scenario for the data argumentation with an aim to balance a dataset and achieved a dataset with higher values, a helpful mechanism in these coarse data situation. Data arbitration steers new image generation for each imbalance image class.

We placed data objects of *FER-2013* dataset into three sets: train, test, and validation set. Over basic sets, objects are placed uneven among emotion labels. Though, with adapted data arbitrations dataset is balanced and aligned for the proposed hypothesis. Figure 5 illustrates both scenarios for data preparation.

The data preparation listed models on images require new sample generation and image normalization to match input shape of 224 × 224 × 3. To ensure consistency across all three models, a standardized input with a batch size of 128 has been employed as the base input. Additionally, a *learning rate* of 0.01 has been set, along with the utilization of L_2 *kernel regularization* and *dropout* techniques to mitigate overfitting risks. Given the objective of classifying multiple emotions, a tailored loss function in the form of *sparse categorical cross-entropy* has been employed.

The experiential works are based on two optimizers: *adaptive momentum estimation (Adam)* and *stochastic gradient descent (SGD. ReLU* has been adapted as activation function, and for classification layer, *softmax* activation is used. *Hyperparameters* are listed in Table 1.

Since *FER-2013* dataset is imbalanced, i.e., dataset contains only 486 '*Disgust*' images and over 7000 images for '*Happy*'; ***data argumentation*** such as ***Width Shift, Height Shift, Shear Range, Zoom, Horizontal Flip, Fill Mode, and Rotation*** used and generated seven new images. For an image with '*Disgust*' emotion with following parameters and deleted images, new images are generated as shown in Fig. 6.

Table 1 **a** List of hyperparameters, **b** data argumentation parameters

Parameter name	Value	Argumentation type	Value
Input shape	224 × 224 × 3	Fill mode	Nearest
Batch size	128	Width shift	0.21
Epoch	100	Height shift	0.2
Learning rate	0.01	Shear	0.2
Optimizer	Adam, SGD	Zoom	0.21
Kernel regularizes	l2(0.01)	Horizontal flip	True
Loss	Sparse CC entropy	Rotation	45°
Activation	Classification layer (*Softamax*), *ReLU*		
Dropout (%)	20, 50		

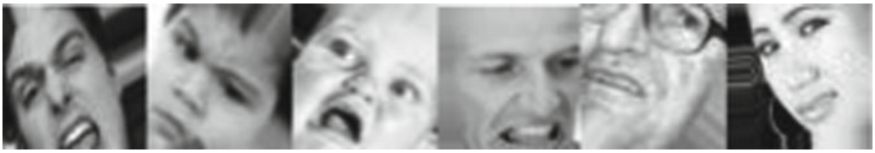


Fig. 6 Data argumentation-based generated images for '*Disgust*' emotion

4.3 Accuracy Evaluation of Fusion-Based Emotion Recognition

In facial emotion work, three models are used as emotion classifier. The key performance criteria observed are its effectiveness on the estimation of emotion. The measure precision is adapted with a definition on Eq. 1. The proposed model precision of formalized as its capacity on the precisely identifying the emotions and their labels:

$$P = \text{True Positive}/(\text{True Positive} + \text{False Positive}). \tag{1}$$

The measure is *recall* and adapted definition is

$$R = \text{True Positive}/\text{True Positive} + \text{False Negative}. \tag{2}$$

The measure of $F1$-score is harmonic mean of both *precision* and *recall* values for a model. The $F1$-score indicates the overall accuracy of a system and formalized as Eq. 3, as

$$F1 = 2 * (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}). \tag{3}$$

The performance analysis reports of listed models are in the form of classification report. Further classification report is placed into three measures: *precision, recall, and F1-score*. The accuracy on correctly outlining the fundamental emotions, these levels on accuracy are illustrated over '*confusion metric'*, depicts the overall accuracy of the system, as for each emotion we can estimate and analyze.

The effective analysis over the designed models is primarily based on fundamental measures and traditional definition. As seen in *confusion metric, the* most correctly classified emotion is '*Happy*' and most less recognized emotion is '*Disgust*'. In case of '*Disgust*', mostly it is misclassified with '*Angry*'; as seen in below image, it is even difficult for human to recognize that it is '*Angry*' or '*Disgust*' (Fig. 7).

'*Neutral*' has also been misclassified with '*Sad*'; as we see in images, it is difficult to tell whether a person is '*Sad'* or '*Neutral*'. Other potential cause for misclassification will be the dense network of *VGG-1* that causes an overfitting and due to addition of our architecture at the end additional parameters has been increased to

Fig. 7 Schematic view of human emotion impressions

139 M which is the major cause of overfitting because amount of data is less for huge number of parameters.

4.4 Results

Modified MobileNetV2 has required around 100 h of training *to converge and training for the defined framework on facial features, with training every* epoch *on validation test*. After the completion of the training phase, model is tested upon test set correct validation of recognition rate. For evaluation purpose, precision, recall, and $F1$-score are used, as shown in Fig. 8.

The overall accuracy boiled down to 64.2% for test set and for validation set accuracy is around 63.9%. Confusion matrix for the same is shown in which it is observed that happy emotion is the most correctly recognized emotion.

Modified ResNet50 It has required around 135 h of training so that model converges completely. During the training phase, model is tested after every epoch on validation test, and after the completion of the training phase, model is tested upon test set correct validation of recognition rate. For evaluation purpose, *precision, recall, and F1-score* are used.

The overall accuracy boiled down to 59.4% for test set and for validation set accuracy is around 58.7%. Confusion matrix for the same is shown in which it is

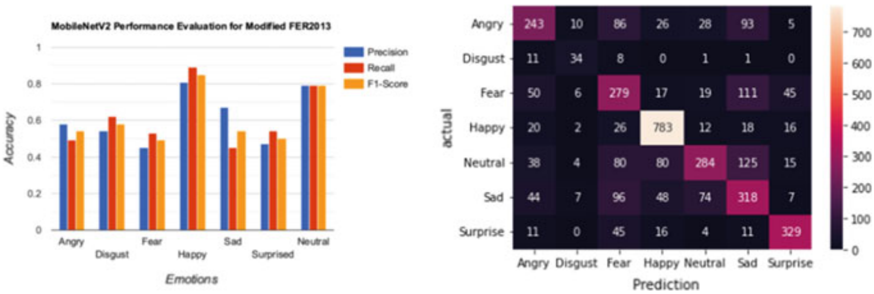


Fig. 8 Performance evaluation chart for *MobileNetV2*

Fig. 9 Performance evaluation and chart for *ResNet50*

observed that happy emotion is the most correctly recognized emotion, as shown in Fig. 9.

Modified VGG-16 It has required around 170 h of training so that model converges completely. *VGG* architecture is very dense in nature due to which it requires more time as compared to other models. During the training phase, model is tested after every epoch on validation test, and after the completion of the training phase, model is tested upon test set correct validation of recognition rate. For evaluation purpose, *precision, recall, and F1-score* are used, as shown in Fig. 10.

The overall accuracy boiled down to 53.5% for test set and for validation set accuracy is around 52.9%. Confusion matrix for the same is shown and can be observed that happy emotion is the most correctly recognized emotion.

Figure 11 illustrates a comparison among all the leaning models over facial features for emotion recognition, here *MobileNetV2* outperforms both the models, i.e., *ResNet50* and *VGG-16,* for the recognition task over all the emotion labels.

Further, comprehensive comparison is conducted upon modified models and other existing models. The redesigned model indicated overall improved performance and outperforms some of the existing models, e.g., *Fast R-CNN*, as shown in Table 2.

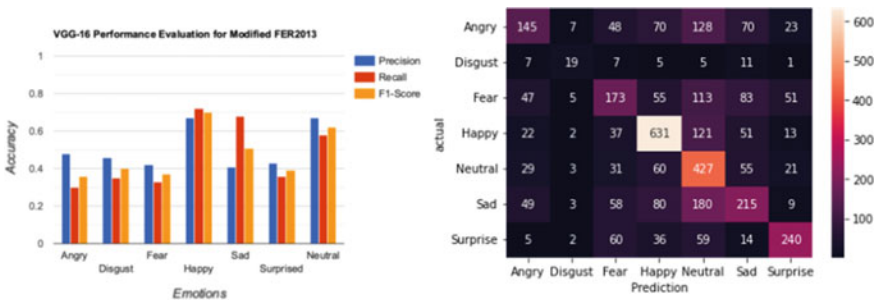


Fig. 10 Performance evaluation chart for modified *VGG-16*

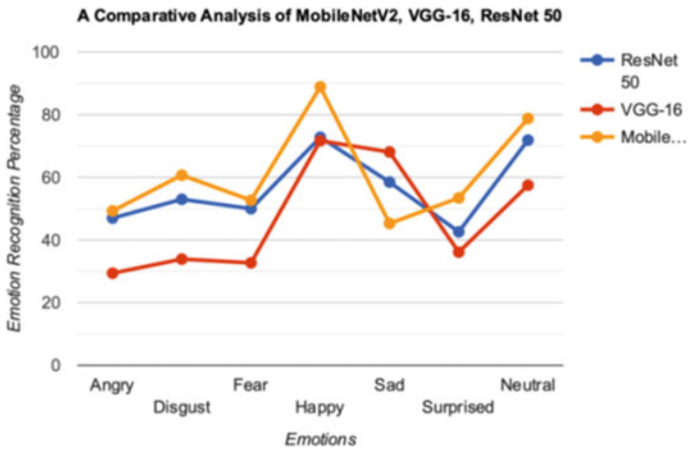


Fig. 11 Comparative analysis of modified model, i.e., *MobileNetV2, VGG-16, ResNet50*

Table 2 Comparison of recognition rate

S. No.	Learning models	Recognition rate (%)
1	**Modified MobileNetV2**	**64.2**
2	**Modified ResNet50**	**59.4**
3	**Modified VGG-16**	**53.5**
4	CNN (AlexNet)[67]	61.1
5	Net B[66]	60.91
6	Net B_DAL[66]	58.33
7	Net B_DAL_MSE[66]	58.15
8	Fast R-CNN(VGG-16)	30.19

The speech-based features play pivotal role on the recognition of emotion in proposed framework, and the models are trained over *SAVEE* dataset and fused with facial feature-based model.

The designed fusion-based emotion recognition framework is firmly placed with two potential equivalents for analysis: Yongqiang Li summation method-based recognition approach and traditional concatenation layer-based method. Our work is conceptually aligned to concatenation-based strategy with summation placed within internal layers.

Figure 12 outlines a relative outcome of recognition of seven human emotions on three layers; it is evident that proposed model performs quite well on few emotions, i.e., for *Disgust* emotion recognition rate increased by 1.8%, *Happy* emotion classification rate increased by 1.3%, of *Surprise* emotion classification rate increased by 6.4%, etc. The evaluation work asserts that with improved trade-offs overall accuracy be significantly enhanced from 90.32% to 91.25%.

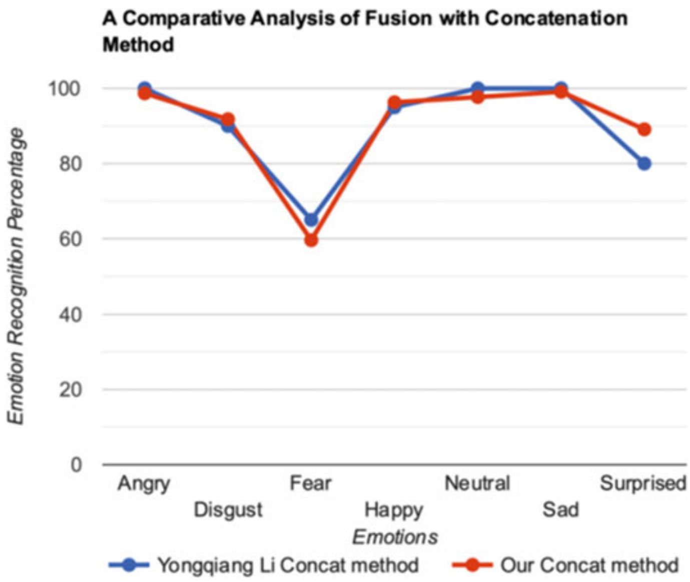


Fig. 12 Comparison of proposed *fusion-based* framework with equivalents

The state-of-the-art observations potentially identified that in all three approaches, the correctly classified emotion is *Happy*, primarily due to availability of massive images and speech samples, whereas the most incorrectly classified emotion is *Disgust* due to lesser number of samples available to train the computation model. Our framework classifies '*Disgust*' emotion with improved rate to some extent.

5 Conclusion

This work aims to design the fusion-based approach for human emotion recognition, with an objective to accurately identify seven human emotions using facial and speech features. The facial domain learning models are based on *FER-2013* dataset with custom layers embedded into predefined model. Here the proposed work surpassed to a *Fast R-CNN* with an increase of accuracy around 3.1% with our addition layers in *MobileNetV2*, wherein speech features are adapted on two different techniques, speech spectrograms which have been extracted from *RAVDESS* as well from combination of *RAVDESS, TESS*, and *SAVEE, and* after extraction process is completed, these spectrograms are fed to our deep learning models, with an increase of 1.3% in recognition rate.

In proposed fusion-based model, at concatenation layer, a novel created model is structured in previous sections. During fusion level, *SAVEE* dataset separated the audio and video and then converted video signals into images, where intensity of emotion is high and then fed all the data to neural network. Interesting thing about this technique is that it takes speech data as well as image data as their input and gives the prediction.

Following potential future research directions are observed:

(i) Data argumentation is one of the key preprocessing activities within a learning-based model. During conceptual design, it has been widely realized that due to the scarcity of data in *FER-2013* dataset, the *SMOTE* could be utilized to generate synthetic data samples.
(ii) In the proposed approach, emotion recognition is delivered over seven human emotions based on two modalities (facial and speech) with an accuracy of 90.32%: though the model is unable to scale up its potential for additional human emotions. In the future solutions, additional modalities such as text, gestures, and physiological signals may be adapted for the recognition of these emotions.
(iii) Current proposed work may be extended to accurately estimate the intensity levels within a specific human emotion to highlight the secondary emotions; e.g., for *Happy* emotion intensity levels could be *Ecstatic*, *Serenity,* and *Joy*.
(iv) One of the key observations from the work is that the current need of scalable fusion classifier recognizes emotions regardless of age, gender, group, ethnicity, stance, lighting, and hair styles.
(v) For a speech-based emotion recognition approach, *delta features* may be employed along with the features *mfcc*, *mels* to strengthen the role speech features in recognition task.
(vi) Dynamic filtering can be used while training models; i.e., one can define their own filters and use a dynamic approach so that filters will change accordingly.
(vii) At last, different combination of fusion layers, i.e., average, weighted sum, weighted average sum, etc., may be experimented in futuristic fusion-based learning to enhance the overall performance.

References

1. Akhand MAH, Roy S, Siddique N, Kamal MA, Shimamura T (2021) Facial emotion recognition using transfer learning in the deep CNN. Electronics 10(9):036
2. Sinha A, Aneesh RP (2019) Real time facial emotion recognition using deep learning. Int J Innov Imple Eng 1
3. Mollahosseini A, Chan D, Mahoor MH (2016) Going deeper in facial expression recognition using deep neural networks. In: 2016 IEEE winter conference on applications of computer vision (WACV). IEEE, pp 1–10
4. Keskar NS, Socher R (2017) Improving generalization performance by switching from adam to sgd. arXiv preprint arXiv:1712.07628
5. Anas H, Rehman B, Ong WH (2020) Deep convolutional neural network based facial expression recognition in the wild. arXiv preprint arXiv:2010.01301
6. Khanzada A, Bai C, Celepcikay FT (2020) Facial expression recognition with deep learning. arXiv preprint arXiv:2004.11823

7. Ivanovsky L, Khryashchev V, Lebedev A, Kosterin I (2017) Facial expression recognition algorithm based on deep convolution neural network. In: 2017 21st conference of open innovations association (FRUCT. IEEE), pp 141–147
8. Gupta R, Vishwamitra LK (2021) Facial expression recognition from videos using CNN and feature aggregation. Materials Today: Proceedings
9. Minaee S, Minaei M, Abdolrashidi A (2021) Deep-emotion: facial expression recognition using attentional convolutional network. Sensors 21(9):3046
10. Barsoum E, Zhang C, Ferrer CC, Zhang Z (2016) Training deep networks for facial expression recognition with crowd-sourced label distribution. In: Proceedings of the 18th ACM international conference on multimodal interaction, pp 279–283
11. Liu W, Zheng W, Lu B (2016) Emotion recognition using multimodal deep learning. Neural information processing, pp 521–529
12. Akçay MB, Og˘uz K (2020) Speech emotion recognition: emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers. Speech Commun 166:56–76
13. Zheng WQ, Yu JS, Zou YX (2015) An experimental study of speech emotion recognition based on deep convolutional neural networks. In: Affective computing and intelligent interaction (ACII), 2015 International conference on. IEEE, pp 827–831
14. Trigeorgis G, Ringeval F, Brueckner R, Marchi E, Nicolaou MA, Schuller B, Zafeiriou S (2016) Adieu features? end-to-end speech emotion recognition using a deep convolutional recurrent network. In: Acoustics, speech and signal processing (ICASSP), 2016 IEEE international conference. IEEE, 5200–5204
15. Ranganathan H, Chakraborty S, Panchanathan S (2016) Multimodal emotion recognition using deep learning architectures. In: 2016 IEEE winter conference on applications of computer vision (WACV). IEEE, pp 1–9
16. Tsironi E, Barros P, Weber C, Wermter S (2017) An analysis of convolutional long short-term memory recurrent neural networks for gesture recognition. Neurocomputing 268:76–86

Concept for Using 5G as Communication Backbone for Safe Drone Operation in Smart Cities

Stefan Kunze, Bidyut Saha, and Alexander Weinberger

Abstract Civilian drones have a wide range of applications and will be an integral part of future smart city designs. While drones fly mostly unregulated today, there will be a strong need for regulation of the lower airspace in the near future. As part of the research project SIMULU, a prototypical geo-awareness system is implemented. This system provides transponder functionality for the drones and is capable of detecting potentially dangerous situations. It can warn pilots and provide them with guidance. For autonomous drones, the autopilot can be updated. In this paper, a theoretical consideration and a concept for using 5G communication as connecting link between drone applications and smart cities are proposed. The possibilities of using 5G as all-in-one communication system, for inter-drone communication as well as enhancing the current geo-awareness system are shown.

Keywords UAV · UTM · 5G · Smart cities

1 Introduction

The basic idea of smart cities is to bring intelligence into urban life, thus increasing the comfort and security of civilians. The concept aims to highly integrate modern IT, like IoT, AI, etc., into urban planning [1]. One aspect of smart city design is the safe integration of civilian (UAS) into the lower airspace. While UAVs have a wide range of applications, their operation remains mostly unregulated, so far. To ensure safe operation in smart cities, this issue has to be addressed by introducing an UTM system. As part of the SIMULU research project, a prototypical GAS is being implemented. This system provides transponder functionality for the drones and is capable of detecting potentially dangerous situations. It can warn pilots and provide them with guidance. For autonomous drones, the autopilot can be updated.

S. Kunze (✉) · B. Saha · A. Weinberger
Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung 94078, Germany
e-mail: stefan.kunze@th-deg.de
URL: http://www.dit.edu

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_10

The new 5G standard is an important backbone for the demanding communication requirements in smart cities. In this paper, a concept for using 5G as a combining link to safely and effectively integrate UAVs into smart cities' lower airspace is proposed. In the following section, a brief overview of some related work is presented. In Sect. 3, the geo-awareness system in its current form is briefly introduced. In Sect. 4, three ways for improving UAS applications in smart cities by introducing 5G are presented. Finally, the paper is concluded with a brief look at some future work.

2 Related Work

Using dedicated radio links for UAVs has a number of disadvantages, such as scarcity of available spectrum, costs associated with separate infrastructure, and incompatibility between systems or connectivity in (BLOS) scenarios [2]. Therefore, the use of cellular networks is discussed by several authors, such as Baltaci et al. [3] (comprehensive overview of drone communication) or Kukliński et al. [4], who propose a detect, sense, and avoid system.

With 5G, a new generation of cellular network has been introduced recently. Compared with 4G/LTE networks, it provides some advantages which are relevant, both for smart city and UAS applications. Besides the commonly known high data rates and lower latencies, 5G also offers the possibility of network slicing [4]. By dividing one physical network into multiple virtual networks, it can provide best possible QoS for a wide range of applications. For UAS applications, 5G provides better air coverage, due to mMIMO beamforming [5]. It can fulfill the communication requirements of UAVs, in terms of data rate and latency and can support moving targets with speeds of up to 500 km/h [6].

A general system architecture for realizing 5G-based (UAS) communication is proposed in [6]. The authors present three different application modes. In "common network sharing" mode, UAVs and other 5G terminals use the same physical and logical network, whereas in the "common network private mode", they use separate logical networks. Finally, in "dedicated private network mode", they are in completely separated physical and logical networks. Si-Mohammed et al. propose a novel cellular network-based architecture for an end-to-end (UAV) business process within the scope of the European Union's U-space [7]. Three major components of this architecture are customer (end user and business provider), U-space (UAV) operator, 5G network owner, (UTM, etc.), and 5G infrastructure (gNodeB). Besides U-space, there are various other approaches for a reliable UAS, as it is clear that the lower airspace needs regulation, surveillance, and control in order to cope with the increasing number of UAVs. Several aviation safety agencies, such as FAA and EASA, presented concepts for UTM [8]. Major UTM functions include UAV registration, UAV database management, UAV flight path management, and continuous monitoring of UAVs during the flight progress [9]. Thus, an integral part of any (UTM) approach will be UAV-borne transponders, which regularly transmit the drone's position and other relevant information to the UTM. Even though the feasibility of ADS-B for drone applications has been demonstrated, the increasing number of UAVs could saturate the ADS-B spectrum [3]. Therefore, it is not a suitable technology for UAS applications in smart cities. An ADS-B like communication using 4G/LTE, LoRa, or XBee is proposed in [10]. First commercial off-the-shelf (UAV) transponders based on LTE cellular networks have recently become available (e.g., Droniq HOD4TRACK or Aerobits The HOD).

3 SIMULU Geo-Awareness System

As part of the SIMULU project, a prototypical geo-awareness system is implemented. It is based on the previously proposed concept [11]. The system allows to warn the pilot of manually guided drones and provide guidance for safely handling dangerous situations. In case of autopilot-guided UAVs, the configuration of the autopilot can be changed and updated in order to restore safe flying conditions. For this purpose, each drone is equipped with an UAV adapter. This device generates transponder messages (containing information like (GNSS) coordinates, altitude, speed, or waypoints if applicable) and regularly transmits them to the SIMULU central system. There, the transponder messages of all connected drones are analyzed. When a potentially dangerous situation (e.g., (UAV) in a no-fly zone or in the corridor of a registered flight) is detected, the central system can either update the autopilot configuration of autonomous drones or warn and provide guidance for human pilots. Warnings and guidance are displayed on the (UI).

For the communication, off-the-shelf radio modules operating in the 868 MHz (ISM) band are used. The communication scheme for a manually piloted drone is illustrated in Fig. 1. The average data rate required for transmission of transponder messages with an update interval of one second is approximately 6.4 kbps, which is at the lower end of the range required for transmission of telemetry data (5–150 kbps, given by Baltaci et. al [3]). For the current prototypical implementation, the chosen radio modules and the dedicated radio channel fulfill all requirements and provide high flexibility. A more performant communication system (let alone 5G) is not needed. For a commercial smart city application, however, this approach

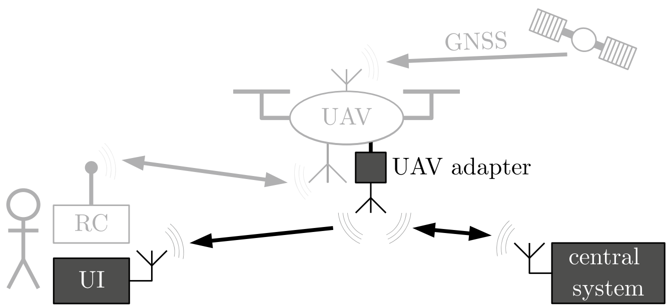


Fig. 1 Current implementation of the geo-awareness system, based on [11]

is not feasible. The ISM band could easily get congested with increasing numbers of UAVs. Also building an area-wide infrastructure dedicated just for this application seems quite impracticable, considering the existence of public cellular networks.

4 Concept

Smart cities deeply integrate modern (IT) components, such as (IoT), (AI), cloud computing, and big data. The overall objective is having "information at your fingertips". As an upcoming new communication technology, 5G has some basic characteristics, such as lower energy, lower cost, higher security, higher reliability, and improved transmission efficiency [12], which make it suitable as smart city communication network backbone. During the design of future smart cities, 5G network planning and urban planning can be merged to provide adequate connectivity for a wide range of communication applications with highly diverse requirements. In Fig. 2, the three main scenarios of 5G and some example applications relevant for smart cities are illustrated. The eMBB provides high data rates required for applications like VR or AR, whereas mMTC targets information exchange between very large numbers of machines, sensors, and other IoT devices. Finally, URLLC provides low-latency communication (down to 1 ms) in combination with high reliability. Both aspects are required for applications like autonomous driving or industrial automation.

Modern civilian drones can be used for a wide range of applications. With the introduction of smart cities, their numbers in the sky will continue to increase. They can be used for autonomous deliveries, or as sensor carrying platforms for inspection and surveillance tasks. There are also applications in law enforcement or emergency services. In the following three subsections, different ways of using 5G as the connecting link between drones and smart cities are presented.

4.1 5G as All-In-One Communication System

Many UAVs require other communication channels (e.g., data streaming link, telemetry link, etc.) besides the (C2)-link. Often multiple different communication systems

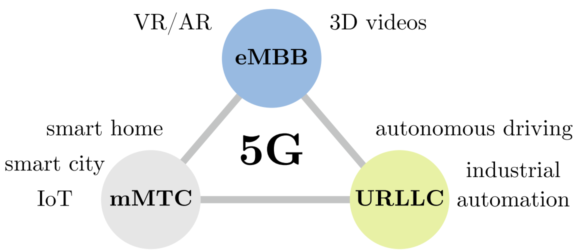


Fig. 2 5G application scenarios for smart city, based on [12]

are used by the same UAS, as each communication system is suitable for a specific requirement. An overview of many different systems used for UAS is given by Baltaci et al. [3]. The introduction of UTM and the geo-awareness system also introduces the need for further communication. Different UAV applications demand different communication needs in terms of latency, data rate, or range. For example, the (C2)-link demands a low-latency communication, whereas 4K video streaming demands a high data rate. Comparing the demands with Fig. 2, it is clear that eMBB and URLLC are also use cases for drones. With the increasing number of drones in smart cities, mMTC also becomes relevant. Neither of the communication systems commonly used for drones is capable of handling all these scenarios. In addition, some of the currently used or proposed communication systems may work well in rural areas, but these are not suitable for complex smart city scenarios. For example, congestion of ISM or ADS-B frequency bands would be probable. As shown in Fig. 2, 5G can meet all those communication requirements required for (UAV) applications. The eMBB is capable of handling the high data rate required for the streaming of sensor data, URLLC satisfies latency and safety requirements for (C2)-links, telemetry, and (UTM) communication, and lastly, mMTC allows to cope with an enormous increase of autonomous UAS.

Considering RF spectrum is a scarce resource in smart cities, it would be advantageous to unify all drone communication into a single communication system. Therefore, the use of 5G as all-in-one communication system for UAS applications is proposed. As illustrated in Fig. 3, it is suggested to include C2, telemetry, payload (video or sensor data), and all (UTM)-related communication, such as transponder messages, warnings/guidance for human pilots, or commands (e.g., autopilot updates) within a common communication channel. This allows to simplify drone design by eliminating the need for various parallel radio systems and helps to use the available spectrum more efficiently. Being a cellular network with the possibility to handover between base stations, 5G is also very suitable to provide BLOS coverage. As the GCS and UAV do not necessarily need to be within the same cell, less transmit power is required for long-range operation, thus making the communication more efficient, both from energy and spectral (less interference/less congested spectrum)

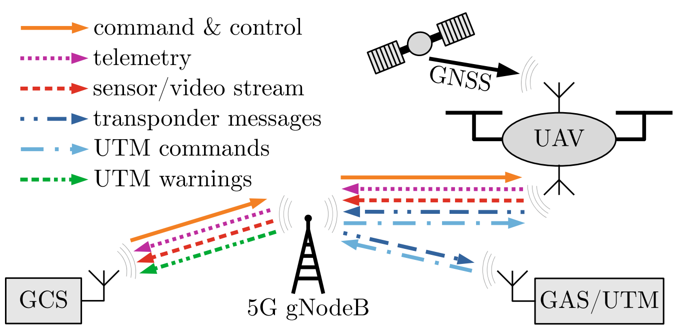


Fig. 3 "All-in-one" concept for 5G integration with SIMULU

points of view. UTM requires the connection to a central entity. For the UTM, an infrastructure with a reliable area-wide coverage is required. Instead of building a new system in parallel, it is more practical and economical to use an existing cellular network. The existing 4G infrastructure can fulfill most of the all-in-one concept's requirements but has some limitations (e.g., for low latency or very high data rate applications). Additionally, 5G introduces the possibility of network slicing, which can segment a single physical network into various virtual networks. This provides the opportunity to configure multiple network slices differently, depending on the communication requirements (e.g., optimized for high data rates or very low latency). It also enables the possibility to create a separate network slice (i.e., virtual network) dedicated to UAS, which could increase security for UTM traffic. Overall, 5G is the most suitable technology for adopting an all-in-one communication concept for UAS.

4.2 Inter-UAV Communication

5G is also a suitable choice for inter-(UAV) communication, which is required for coordination of drone swarms or collision avoidance. A concept discussing the feasibility of 5G-based drone swarms is presented in [13]. With the introduction of (UTM) as central point, there are two possible ways for 5G-based inter-drone communication (see Fig. 4). In the first approach, all possible communication is sent through UTM system (i.e., (UAV) 1 sends a message to (UTM) which relays it to (UAV) 2). In the second approach, UAVs can communicate with each other directly (via the gNodeB), without the UTM.

In this paper, a mixture of those two methods is proposed for inter-(UAV) communication. The goal of the UTM is to have an overview of all (UAV) movements and coordinate them to assure safe operation. Thus, all transponder messages from drones must always be sent to the (UTM). There the overall (UAS) traffic is monitored, and information from UAV x can be relayed to UAV y if necessary, to coordinate their flight paths. This approach may be applicable for collision detection, as the UTM has the "bigger picture" and is capable of recognizing potential collisions. The second approach is suitable for cases, where the content of communication is not (directly) relevant for the UTM. This is the case for coordination of drone swarms, as each

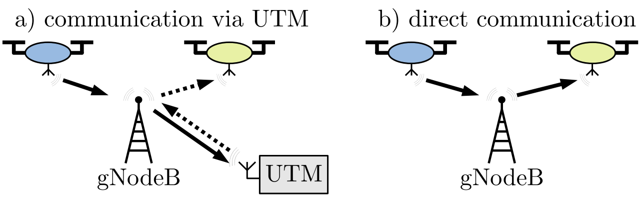


Fig. 4 Possible inter-UAV communication strategies

individual UAV still sends its transponder messages. By removing the UTM from the communication path, the latency can be reduced, which is beneficial for time-critical URLLC traffic. A second application which can also benefit from reduced latency may be collision avoidance. Once the threat of an imminent collision is detected, two UAVs can directly negotiate evasive maneuvers without involving the UTM. This data exchange may contain more detailed data (e.g., more detailed telemetry) at a shorter interval than the one which is used for regular transponder messages sent to the UTM.

4.3 Improved Geo-Awareness System

Using 5G for the geo-awareness system's communication instead of a dedicated radio link offers several advantages. In future smart cities, a reliable 5G coverage can be taken for granted. Therefore, using 5G would avert the need (and subsequent costs) of building a new infrastructure just for geo-awareness system and UTM communication, while at the same time, taking a step toward using the already scarce frequency spectrum more efficiently.

With the current geo-awareness system's implementation, only UAVs that are equipped with an autopilot can be influenced directly by updating the mission (e.g., assigning new waypoints). For manually controlled drones, only warning and guidance for the pilot are possible. By using 5G for the UAS communication, enough data rate and sufficiently low latencies can be assured, to actively let the geo-awareness system perform maneuvers with such drones. While there are certainly limitations to this approach for more complex flight maneuvers, simple ones (e.g., changing altitude) may in many cases be sufficient to avoid dangerous situations. When the drone is actively controlled by the GAS or when multiple UAVs operate in close proximity, it may be necessary to provide more detailed or more frequent transponder messages. Compared to the currently used ISM-band radio modules or other potential systems like LoRa, 5G offers more performance reserves in this regard. Since 5G is also specified for (mMTC), it is capable of handling the expected increase in UAS flight movements in smart cites. For the geo-awareness system's user interface, the introduction of 5G offers two advantages. On the one hand, any 5G capable smartphone or tablet can be used to replace the current (UI). On the other hand, more advanced and complex guidance mechanisms (like (VR)/(AR)) can be realized to assist human drone pilots. The ability of 5G to create multiple virtual networks within one cell allows to optimize the network for multiple applications, by configuring the network slices accordingly. Additionally, safety and security relevant traffic (such as (UTM) communication) can be separated from "normal" cellphones. The possibility of broadcasting messages within one network slice allows to efficiently propagate information to all UAVs within a cell. Overall, the combination of the current prototypical geo-awareness system and the proposed all-in-one communication concept provides improvements to make the system more intelligent, more generic (no need for own infrastructure), and more cost effective, thus making it a valuable asset for improving the safety in the lower airspace of smart cities.

5 Conclusion and Future Work

In this paper, a concept for using 5G as connecting link between drones and smart cities is proposed. Using 5G seems to be the best-suited communication system for the presented use cases. In the next steps, the proposed mechanisms (like all-in-one communication with multiple network slices, 5G-based takeover of manually piloted drones, etc.) have to be implemented to proof their feasibility, before the overall concept may be implemented for smart cities.

Acknowledgements The work presented in this paper is part of the SIMULU project, which is funded by the German Federal Ministry for Digital and Transport.

References

1. Kumar NM, Goel S, Mallick PK (2018) In: 2018 Technologies for smart-city energy security and power (ICSESP). https://doi.org/10.1109/ICSESP.2018.8376669
2. Goddemeier N, Daniel K, Wietfeld C (2010) In: 2010 IEEE Globecom workshops, pp 1760–1765. ISSN 2166-0077. https://doi.org/10.1109/GLOCOMW.2010.5700244
3. Baltaci A, Dinc E, Ozger M, Alabbasi A, Cavdar C, Schupke D (2021) IEEE Commun Surveys Tutor 23(4):2833. https://doi.org/10.1109/COMST.2021.3103044 (Conference name: IEEE Communications Surveys & Tutorials)
4. Kukliński S, Tomaszewski L, Korzec P, Kolakowski R (2020) In: 2020 6th IEEE conference on network softwarization (NetSoft), pp 242–246. https://doi.org/10.1109/NetSoft48620.2020.9165458
5. Bhuyan A, Guvenc I, Dai H, Yapici Y, Rahmati A, Maeng SJ (2019) In: 2019 IEEE 90th vehicular technology conference (VTC2019-Fall), pp 1–5. ISSN 2577-2465. https://doi.org/10.1109/VTCFall.2019.8891595
6. Yan K, Ma L, Zhang Y (2020) In: 2020 IEEE 9th joint international information technology and artificial intelligence conference (ITAIC), vol 9, pp 1115–1118. ISSN 2693-2865. https://doi.org/10.1109/ITAIC49862.2020.9339133
7. Si-Mohammed S, Bouaziz M, Hellaoui H, Bekkouche O, Ksentini A, Taleb T, Tomaszewski L, Lutz T, Srinivasan G, Jarvet T, Montowtt P (2021) IEEE Veh Technol Mag 16(1):57. https://doi.org/10.1109/MVT.2020.3036374 (Conference name: IEEE Vehicular Technology Magazine)
8. Bekkouche O, Bagaa M, Taleb T (2019) In: 2019 IEEE global communications conference (GLOBECOM), pp 1–6. ISSN 2576-6813. https://doi.org/10.1109/GLOBECOM38437.2019.9014200
9. Park JH, Choi SC, Ahn IY (2019) In: 2019 Eleventh international conference on ubiquitous and future networks (ICUFN), pp 118–120. ISSN 2165-8536. https://doi.org/10.1109/ICUFN.2019.8806075
10. Lin CE, Hsieh CS, Li CC, Shao PC, Lin YH, Yeh YC (2019) In: Integrated communications. Navigation and surveillance conference (ICNS), pp 1–12. ISSN 2155-4951. https://doi.org/10.1109/ICNSURV.2019.8735350

11. Kunze S, Weinberger A (2021) In: 2021 31st International conference Radioelektronika (RADIOELEKTRONIKA), pp 1–6. https://doi.org/10.1109/RADIOELEKTRONIKA52220.2021.9420196
12. Chen H, Yuan L, Jing G (2020) In: 2020 2nd International conference on artificial intelligence and advanced manufacture (AIAM), pp 154–157. https://doi.org/10.1109/AIAM50918.2020.00038
13. Campion M, Ranganathan P, Faruque S (2018) In: 2018 IEEE international conference on electro/information technology (EIT), pp 0903–0908. ISSN2154-0373. https://doi.org/10.1109/EIT.2018.8500274

5G Stand-Alone Test Bed for Craft Businesses and Small or Medium-Sized Enterprises

Siegfried Roedel, Frantisek Kobzik, Markus Peterhansl, Rainer Poeschl, and Stefan Kunze

Abstract The switch from 4G to 5G mobile communication leads to a significant increase in opportunities for businesses. 5G not only offers a wider bandwidth connection, but also forms the basis for new applications, business models and products. Thus, it is important for companies to take 5G into account in their developments as soon as possible. In this paper, the concept and implementation of a 5G stand-alone test bed and some exemplary use cases relevant for small businesses are presented. The test bed will serve as basis for developing and showcasing customized 5G applications. In combination with other education and training offers, this test bed will provide valuable knowledge transfer to small businesses and help them with the integration of 5G.

Keywords 5G · Mobile communication · Test bed

1 Introduction

With the trend toward interconnected production and value chains as well as IoT, Industry 4.0, cloud computing and AI, the communication demands rise above the capabilities of LTE advanced and Wi-Fi. 5G is the first generation of cellular networks, which can be widely used for industrial processes. Some reasons for this are performance parameters in terms of bandwidth, latency and reliability. The easy implementation of isolated campus networks and the possibility of slicing the network into multiple virtual networks (each optimized for a specific application) are also important advantages of 5G. All these factors lead to the increasing popularity of 5G in the industry.

SMEs and even small craft businesses could also benefit from 5G (e.g., by enabling new business models, integration in products, or having a customized cellular network under their own management). However, several challenges and obstacles are

S. Roedel (✉) · F. Kobzik · M. Peterhansl · R. Poeschl · S. Kunze
Deggendorf Institute of Technology, Institute for Applied Informatics, Freyung 94078, Germany
e-mail: siegfried.roedel@th-deg.de
URL: http://www.dit.edu

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_11

stopping them from implementing their own 5G solutions. The high capital and operational expenses of a 5G campus network can be a problem for such businesses. The lack of knowledge or staff training in regard to wireless communication (cellular networks in general and 5G in particular) is another factor. Many 5G application scenarios in the field of (IIoT) are designed and specified to meet the requirements of big industries or corporations. Hence, there is a lack of market-ready applications that suit the need of smaller businesses. These challenges in combination with the absence of best-practice examples to demonstrate the advantages of 5G for small businesses are the reasons why most of them shy away from implementing 5G.

To address these obstacles and challenges, the Deggendorf Institute of Technology started a new project in cooperation with the Chamber of Crafts in Lower Bavaria and Upper Palatinate. This project looks into the problems by developing and implementing various 5G applications tailored to the needs of craft businesses and SMEs. A structural and technological analysis of the regional business is performed to find the best ways to help and advice them on 5G-related questions. One central aspect of this project is the implementation of a 5G test bed (stand-alone campus network), which serves as a "playground" for the prototypical development, demonstration and showcasing of customized 5G applications. Through workshops and training courses for business owners and employees, their awareness of 5G and its possibilities will be increased. This will also improve the technology transfer between the research institute and small businesses.

In this paper, the concept and implementation of this test bed and some exemplary applications, which will be developed and demonstrated in the course of the project, are presented. In the following section, a brief overview of some related work and various other 5G test bed implementations are given. In the third section, the architecture of the implemented test bed and the available measurement equipment are presented and some current limitations are discussed. In the fourth section, various exemplary 5G applications which are relevant for craft trades are shown. Additionally, first demonstrators using the test bed are discussed. Finally, the paper is concluded with a look ahead at some future work.

2 Related Work

Mobile communication of the fifth generation has great potential for various applications. For example, 5G offers the possibility of private networks in the licensed spectrum for the first time. They are the key point why 5G is suited for industrial wireless networking as they offer dedicated coverage, intrinsic control, exclusive capacity, customized service and reliable communication [1].

Despite these advantages, especially SMEs have problems recognizing their specific opportunities. Without realizing the full potential of 5G, many of them shy away from the investment, as the initial costs seem to be too high. Furthermore, 5G is much more complex than Wi-Fi in regard to the software, the necessary adaption of the business organization, and the operation of the network. These reasons lead to a slow

5G adoption, particularly in SMEs and craft businesses. However, the slow adoption could result in reduced competitiveness of the respective companies in the future [2]. To overcome these challenges, exemplary cross-company 5G networks can be very helpful. The 5G introduction in companies (especially SMEs) should also be supported by advanced training for the workers [2]. A promising approach to foster the introduction and application of 5G is to build lighthouse projects. These projects should develop and show descriptive use cases to demonstrate the advantages of the emerging technology [2].

Recently, some corresponding 5G test beds and demonstration projects with different focuses have been established. The authors in [3] propose mobile test beds for 5G (SA) and (NSA) with standard (RAN) components and the open-source Open5GS[1] 5G core. The test bed focuses on industrial automation with the example of a robot that is mounted on an automated guided vehicle. The research institute Fraunhofer FOKUS established an indoor and outdoor 5G playground in Berlin which is operated by the Open5GCore.[2] Their main focuses are smart city applications as well as automotive and industrial applications [4]. A further example is a concept for a flexible 4G and 5G test bed using OpenAirInterface[3] and open source (MANO). The concept has been proven by implementing its basic functionality. However, there are some performance issues [5]. The authors in [6] propose a reference architecture for a distributed 5G test bed based on a cloud-native network functions virtualization (MANO) approach. Besides these more general approaches, there are also test beds with a specific focus. One example is a specialized 5G test bed for delay measurements in stand-alone and non-stand-alone campus networks with (RAN) components [7].

With the emergence of 5G, some use cases and possible applications of the new technology are being researched. For instance, in a white paper of umlaut AG, several use cases of 5G for Industry 4.0 applications are presented. Among others, track and trace applications as well as automated process control, (AR)/(VR) and autonomous transport applications are mentioned. For all of the presented applications, it is expected that 5G will show a better performance than the alternatives (e.g., 4G and Wi-Fi 4/5/6) [8]. The authors of [9] present, among others, a collection of current 5G use cases and their realization. They further summarize some research gaps. The main gap is the lack of real-world implementations and demonstrations in industrial environments. O'Connell et al. analyze the chances and challenges of 5G in manufacturing environments. Several examples for future applications are shown, like the real-time control of robots, (AR)/(VR) applications to support (predictive) maintenance and training as well as the improved tracking capabilities of goods and products. All of this can significantly improve productivity and take manufacturing to a new level [10]. A further example is the usage of a 5G (NSA) platform for the remote control of an automated guided vehicle. The comparison of the vehicle guide error between 4G and 5G shows a notably better performance of 5G due to the lower

[1] https://www.open5gs.org.

[2] https://www.open5gcore.org.

[3] https://www.openairinterface.org.

latency [11]. Besides the opportunities, some challenges must be taken into account. Examples are the interoperability with existing protocols and issues mainly in the security area [10].

3 Implemented 5G Test Bed

3.1 Goals and Requirements

The test bed was constructed with the mindset to research 5G applications. It is also designed to enable partner facilities, companies and developers to experience a working 5G stand-alone network environment. The test bed should enable them to perform use case tests or to evaluate new products. With a focus on craftsmanship in the (SME) area, the goal of the test bed is to evaluate and demonstrate 5G applications that are tailored to craft businesses. Craftsmen, (SME) workers and developers should be able to experience 5G applications first-hand. This allows them to increase their awareness of digitization in general and 5G in particular. At the same time, reservations against new technologies shall be reduced, for example by clearing up misunderstandings. The benefits of the new communication technology shall be demonstrated within relevant use case scenarios.

To achieve these goals, the test bed should implement most of the features introduced with 5G. The test bed shall be a 5G stand-alone Open-RAN network. It shall be upgradeable and expandable for future applications. Usually, a 5G network can only be operated with equipment from a single manufacturer to ensure system compatibility. To maximize the flexibility of the test bed, the system shall conform to the Open-(RAN) standard [12]. Thus, it allows using any equipment (even from different manufacturers) that complies with the standard. Besides that, the system shall have low operation expenses.

3.2 Implementation

In general, a 5G network can be divided into back-, mid- and fronthaul. The connection between the core and CU is called backhaul and can cover a distance of several hundred kilometers. The connection between the (CU) and the (DU) is called midhaul and can span a distance of dozens of kilometers. Finally, the connections between the (DU) and the RUs are called fronthaul and can have a length of several kilometers. A single core can have several CUs, and each of them can be connected to several DUs. Likewise, each (DU) can have several RUs connected to it. With this architecture, the computing power and latency requirements can be shifted between components to create the desired network flexibly.

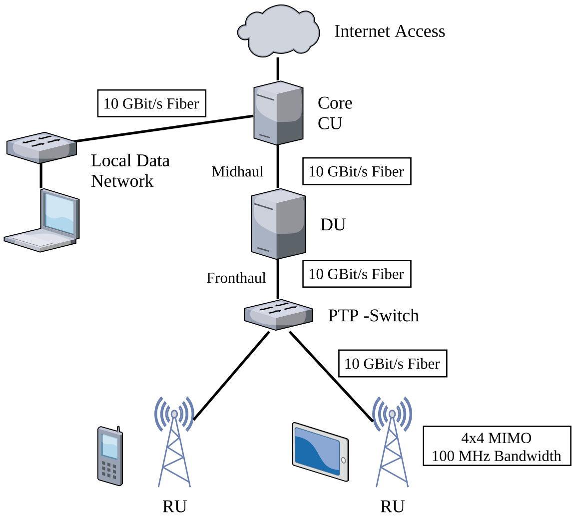


Fig. 1 5G (SA) network at Deggendorf Institute of Technology

In comparison with public 5G networks, or "ready-to-use" campus network solutions, the test bed implemented at the Deggendorf Institute of Technology is highly customizable. For example, frequency bandwidth, subcarrier spacing and time division duplex slot format can be configured to fit the examined use cases. The overall architecture of the test bed is illustrated in Fig. 1. It is an Open-(RAN) 5G (SA) campus network using frequency band n78 (between 3.7 and 3.8 GHz). This is the regulated frequency range usable for private campus networks in Germany [13].

The network consists of a central 5G core and the RAN which uses Open-(RAN) compatible RUs. The core serves as the main 5G management software. It runs on a commercial off the shelf server. The RAN represents the mobile network base station and handles the wireless communication. It is separated into three parts named CU, DU and RU. For easier understanding, CU and DU can be imagined as the base station on the ground and RU as the antenna on a mast. In this test bed, the CU is running as a software solution on the same server as the 5G core. The DU is a software solution on a dedicated commercial off the shelf server with acceleration cards. The RUs are more than simple passive antennas. They also perform a part of the lower-level base station calculations and have the radio frequency transmitting chipset integrated. Each RU can transmit within frequency band n78 with a signal bandwidth of up to 100 MHz and supports features like 4 × 4 (MIMO).

The 5G test bed uses a simple architecture with one 5G core as a network management and authentication system. The core has a logical connection to the (CU). They run in virtual instances on the same server. The CU handles the higher-layer communication and protocol stacks. These protocols include radio resource control or packet data convergence protocol. Both are parts of the layer three network communication between UE and RAN. The lower-layer communication and protocol stacks are handled by the DU, which runs on a separate server. These protocols include radio link control or medium access control. The DU is connected to both

RUs via a (PTP) fiber switch. It uses the global positioning system for highly accurate clock reference and ensures time synchronicity between the connected devices. At the end of the chain, each RU is handling parts of the lower-layer communication and transmits the 5G radio signals.

3.3 Extended Capabilities

In addition to the test bed which is based on closed source components, an open-source-based 5G system running on general-purpose hardware (x86 workstation) is also available. For this system, the open-source Open5GS[4] serves as 5G core. The open-source solution srsRAN[5] is implemented to establish a virtual (gNB), including the (CU) and (DU) functions. To generate the radio signals, a software-defined radio unit is used which is connected to the workstation. To perform safe tests, the software-defined radio is connected to an electromagnetic interference shielding box which prevents unwanted signal disturbances in the public network. Open5GS is compliant with release 16 of the 5G standard and may also be used in combination with some commercial (RAN) systems[4]. With this test setup, it is possible to create a purely open-source-driven 5G network and to compare it with the test bed which is based on a commercial solution.

For an in-depth analysis of 5G applications, some test and measurement equipment is available to supplement our 5G infrastructure. One of these devices is a signal generator that can generate 5G signals with customized transmitting values. This is complemented by a signal analyzer, capable of measuring the signal quality of the 5G transmission. Both devices support frequencies up to 44 GHz. Thus, they may also be used in future mmWave applications.

For comparison measurements between different public 5G networks (operated by multiple carriers), a measurement backpack is available. The backpack allows simultaneous performance and signal measurements for all three major mobile network providers in Germany. For signal measurements, the backpack includes a small signal analyzer that can detect all nearby mobile networks within sub-6 (below 6 GHz) frequencies and part of the mmWave spectrum (between 30 GHz and 300 GHz). Furthermore, three mobile phones are included which can perform several network performance tests such as throughput and latency tests.

3.4 Limitations and Problems

One of the major problems of 5G (SA) networks is finding compatible UE devices. As of September 2022, most of the available devices only support 5G (NSA) networks,

---

[4] https://www.open5gcore.org.

[5] https://www.srsran.com.

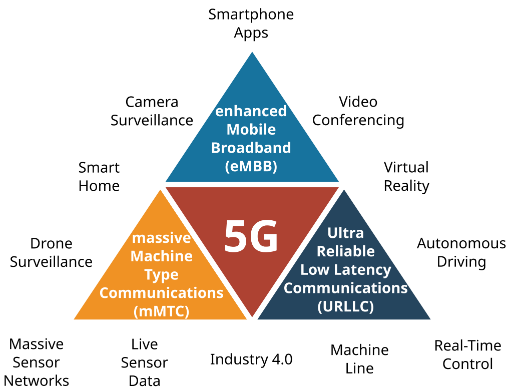


Fig. 2 Application scenarios of 5G

either through hardware or firmware limitations. Therefore, only a couple of devices (mostly 5G routers) could successfully be connected to the 5G (SA) Open-RAN test bed until now. This should be less of an issue with the introduction of new products to the market.

Another problem for private campus networks could be the usable mobile network identification number, as some UEs may not connect to all possible combinations. The identification is usually done via MCC and MNC. MCC declares the country code to which a mobile subscription belongs and MNC the corresponding mobile network operator. In Germany, the default values for private 5G campus networks are 999 for (MCC) and 99 for (MNC), which represent testing networks. The usual MCC for Germany would be 262. To research this potential issue, official German MCC and MNC identification numbers will be applied to the competent authority.

4 Applications

5G allows the implementation of applications that were not feasible with previous generations of mobile networks. It covers three main use case classes: eMBB, uRLLC and mMTC. Each of them provides different connection features (data transfer speed, low latency and reliability and high (UE) density), as shown in Fig. 2. In this section, typical examples for each class are presented. Additionally, the first demonstrators that were built using the test bed and their relevance to craft businesses and SMEs are discussed.

The mMTC allows connecting a vast number of devices to the 5G network (theoretically up to 10^6 devices per square kilometer [14]). Thus, mMTC supports massive IoT-like applications, such as high-density sensor networks (in industry, smart agriculture and chemistry), monitoring and tracking of large numbers of devices (asset tracking in industries), or improvements in company logistics (e.g., tracking and coordinating big numbers of vehicles). The uRLLC class allows implementing applications that have strict requirements on network latency (with mmWave equipment, down to 1 ms and below) and data transfer reliability (>99.9999%) [15]. However, no 5G implementations reaching these goals are known to us yet. The possible applications of uRLLC include real-time control of robots or vehicles (e.g., warehouse logistics, piloting drones), real-time (V2X) communication (driver-less cars and truck platooning), remote maintenance on production plants and machines or (AR)-assisted surgery, where a connection to clean rooms without cables is needed. The eMBB supports applications which need high data transfer rates (up to 20 Gbps in future deployments [16]). This is mainly driven by increasing demand for fast transfer of high-volume data such as video streams to mobile phones. Possible application scenarios of (eMBB) are among others:

- High-volume visual data transfer in the building sector (e.g., inspections of buildings or constructions with drones)
- Live viewing and editing of building information modeling maps on construction sites
- Temporary Internet connectivity for Wi-Fi-based UEs on construction sites
- Permanent Internet connectivity for small businesses which have no sufficient broadband connection (replacement of traditional Internet service providers)
- Streaming high-volume multimedia data in (AR)/(VR)/(MR) devices
- Remote video surveillance.

It is important to mention that the requirements of some applications span multiple use case classes. For instance, (VR)/(AR)/(MR) applications used for controlling machinery require fast and reliable connection provided by uRLLC and the high bandwidth of eMBB. Another example is controlling a large number of business-critical (IIoT) devices, where features of both uRLLC and mMTC are required. Nevertheless, simultaneous usage of multiple network slices within a single network connection is not supported. To partially overcome this problem, a new network slice with parameters tailored to the needs of the application can be created. It is important to note that network slices cannot benefit from all of the mentioned advantages, but would have to use a trade-off. Another possible solution is using separate connections in the application, each utilizing a different network slice. Furthermore, the authors of [17] suggest various sophisticated options of combining uRLLC and mMTC.

To test the implemented 5G stand-alone test bed, demonstrators were developed. They are already focused on possible use cases for craft businesses and SMEs. As a first demonstrator, a simple experiment was performed in the laboratory by streaming a video to virtual reality glasses. The setup consists of the HTC Vive Focus 3 (VR) glasses connected to the test bed's 5G network via the ZTE MU 5001 5G modem. The unbuffered FullHD video with 60 frames per second can be played fluently, even when a handover between two radio units is performed. In the course of the project, this example will be further developed into a demonstrator showcasing the possibilities of an interactive platform for remote staff training in SMEs. The final market-ready application could, for instance, help to train or assist the workers to operate or repair complex machinery in factories in an interactive way. This would significantly reduce

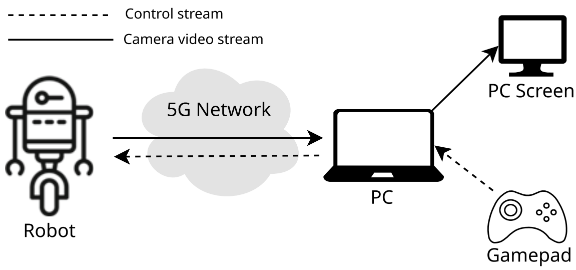


Fig. 3 Robot control and video streaming

the unnecessary traveling of specialized customer service personnel. This results in reducing travel costs as well as downtimes. Additionally, this would also mitigate the risk of damaging the machinery by unprofessional handling.

A second demonstrator was built to show the feasibility of real-time robot control via 5G. The experiment involved the FreeNove Big Hexapod Robot (based on Raspberry Pi 4) equipped with a camera and connected to the 5G test bed using the Quectel RG500Q-GL 5G modem. The 5G connection was operated in the eMBB mode, as uRLLC was not yet supported by the test bed network at that time. The robot exposed two services as illustrated in Fig. 3:

- Live video stream (at 720p resolution) from the camera over TCP
- Control service daemon (based on TCP as well), allowing to move the robot around using a gamepad.

The goal of the experiment is to test the subjective fluency of the video transmission and the remote control responsiveness. The experiment was concluded without noticing any problems, even during handover. The future real-world utilization of this experiment in SMEs could lie in implementing a revision system based on remotely controlled unmanned aerial vehicles. Such systems would simplify and speed up the inspection of objects that are hard or dangerous to access (e.g., constructions or high-rise machines) or vast areas (e.g., forests or fields). Such applications would simplify the operation by multiple users from different locations. This would result in cost savings.

Another 5G application that will be researched in the course of the project is establishing a data connection of a PLC via a 5G router for remote maintenance independent of the company's network (Fig. 4). In big industrial factories, there is usually a maintenance department with experts in different fields (e.g., electronics or mechanics) who can readily repair defective machinery. This is usually not the case in a craft business or (SME), as having a big maintenance department is not affordable for them. Therefore, they have to contact the manufacturer's customer service every time a machine breaks down, which results in increased downtimes. The machine can be connected directly via the 5G router so that the manufacturer's customer service can carry out the fault diagnosis immediately online. Then, remote maintenance or (AR) assisted repairs can be performed without further delay. For craft businesses and SMEs, this approach could result in shorter downtimes and lower maintenance costs.

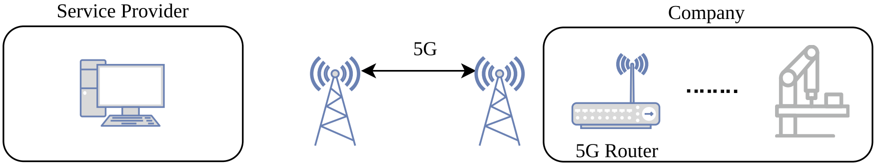


Fig. 4 Management of data connections to PLCs

5 Conclusion and Future Work

Many small companies could benefit from adopting 5G. However, the lack of knowledge and practical experience slow down the adoption of the new technology by craft businesses and SMEs. The implemented 5G test bed is specifically designed for developing and demonstrating use cases that are relevant for these types of companies. This allows them to gather hands-on experience with 5G and helps them with the adoption of 5G into their value chain. The test bed in combination with training courses and seminars also enables a steady transfer of knowledge between the chamber of crafts and the research institute on the one side and small and craft businesses on the other. In the course of the project, the existing demonstrators (like robot control or (AR) applications) will be expanded further. Additionally, new ideas for other possible showcases shall be developed and implemented in close cooperation with the chamber of crafts and interested companies.

New features of 5G (e.g., network slicing or mmWave technology), which will gradually be introduced and implemented in the near future, shall also be taken into account and showcased in relevant scenarios. For example, the mmWave technology enables precise indoor localization services, which could be used for autonomous transport systems. The accuracy will increase from currently about 20 meters in the sub-6 GHz band to one meter using mmWave.

Acknowledgements The work presented in this paper is part of the project "5G für Handwerk und Mittelstand" (5G for craft businesses and small and medium-sized enterprises), which is funded by the Bavarian State Ministry for Economic Affairs, Regional Development and Energy.

References

1. Aijaz A (2020) IEEE Ind Electron Mag 14(4):136
2. Fleischer J, Albers A, Anderl R, Aurich J (eds) (2021) In: 5G in der Industrie - Wege in die Technologieführerschaft in Produktentwicklung und Produktion. acatech IMPULS. acatech, München. https://www.acatech.de/publikation/5g-in-der-industrie/download-pdf/
3. Senk S, Itting SAW, Gabriel J, Lehmann C, Hoeschele T, Fitzek FHP, Reisslein M (2021) In: European wireless 2021, VDE, Verona, Italy, 2021, pp 69–76
4. Fraunhofer FOKUS. NGNI 5G Playground (2022). https://www.fokus.fraunhofer.de/go/en/fokus_testbeds/5g_playground

5. Dreibholz T (2020) In: Web, artificial intelligence and network applications. In: Barolli L, Amato F, Moscato F, Enokido T, Takizawa M (eds) Advances in intelligent systems and computing, vol 1150. Springer, Cham, pp 1143–1153. https://doi.org/10.1007/978-3-030-44038-1_105. http://link.springer.com/10.1007/978-3-030-44038-1_105 (Series title: Advances in intelligent systems and computing
6. Arampatzis D, Apostolakis KC, Margetis G, Stephanidis C, Atxutegi E, Amor M, Di Pietro N, Henriques J, Cordeiro L, Carapinha J, Khalili H, Rehman A (2021) In: 2021 IEEE international Mediterranean conference on communications and networking (MeditCom). IEEE, Athens, Greece, , pp 13–19. https://doi.org/10.1109/MeditCom49071.2021.9647591. https://ieeexplore.ieee.org/document/9647591/
7. Rischke J, Sossalla P, Itting S, Fitzek FHP, Reisslein M (2021) IEEE Access 9:121786
8. Mennig DJ, Hajek L, Münder P (2019) 5G in production. Whitepaper, umlaut AG, Aachen (2019). https://www.umlaut.com/uploads/documents/200331_Whitepaper_5GinProduction_umlaut.pdf
9. Varga P, Peto J, Franko A, Balla D, Haja D, Janky F, Soos G, Ficzere D, Maliosz M, Toka L (2020) Sensors 20(3):828
10. O'Connell E, Moore D, Newe T (2020) Telecom 1(1):48
11. Nakimuli W, Garcia-Reinoso J, Sierra-Garcia JE, Serrano P, Fernandez IQ (2021) IEEE Commun Mag 59(7):14
12. O-RAN Alliance e.V. O-RAN ALLIANCE. https://www.o-ran.org/
13. Bundesnetzagentur. Regionale und lokale Netze.https://www.bundesnetzagentur.de/DE/Fachthemen/Telekommunikation/Frequenzen/OeffentlicheNetze/LokaleNetze/lokalenetze-node.html
14. Chen X, Ng DWK, Yu W, Larsson EG, Al-Dhahir N, Schober R (2021) IEEE J Sel Areas Commun 39(3):615
15. Li Z, Uusitalo MA, Shariatmadari H, Singh B (2018) In: 2018 15th International symposium on wireless communication systems (ISWCS). IEEE, Lisbon, Portugal, pp 1–6. https://doi.org/10.1109/ISWCS.2018.8491078. https://ieeexplore.ieee.org/document/8491078/
16. Osseiran A, Parkvall S, Persson P, Zaidi A, Magnusson S, Balachandran K (2020) 5G wireless access: an overview. In: Whitepaper 1/28423-FGB1010937, Ericsson. https://www.ericsson.com/498a10/assets/local/reports-papers/white-papers/whitepaper-5g-wireless-access.pdf
17. Pokhrel SR, Ding J, Park J, Park OS, Choi J (2020) IEEE Access 8:131796

Cryptography in Latvia: Academic Background Meets Political Objectives

Rihards Balodis and Inara Opmane

Abstract The paper introduces the Cryptography Digital Ecosystem concept and describes the rationale for its implementation. The authors have developed an ecosystem deployment manual and the paper indicates the tasks from the manual for the practical implementation of cryptography solutions in the country. Identified tasks have been analytically developed based on wide analysis of public literature regarding the development of quantum encryption solutions and adequate compliance with the solution requirements.

Keywords Cryptography · Digital ecosystem deployment · Information access regulation · Quantum cryptography · Quantum key distribution (QKD).

1 First Section

1.1 A Subsection Sample

An important research topic in the Centre for Quantum Computing Science of the University of Latvia is quantum computing: the theoretical aspects of quantum in-formation including quantum algorithms, computational complexity, communications, and cryptography.

As solutions of the practical application of quantum computing are in the near future, the strategy of the Institute of Mathematics and Computer Science of the University of Latvia (IMCS UL) is to use quantum technologies that can be applied now and immediately.

The activity of IMCS UL was concentrated in quantum communications and encryption (quantum encryption) applications (taking into account the institute's

R. Balodis (✉) · I. Opmane
Institute of Mathematics and Computer Science, University of Latvia, Raina Bulv. 29, Riga 1459, Latvia
e-mail: imcs@lumii.lv

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_12

previous experience in the introduction of the internet in Latvia and the partnership with GEANT since 2000).

IMCS UL started the development of quantum cryptography research topics in Latvia in 2019. The works were initiated with the purchase and operational testing of Clavis 3 from ID Quantique (www.idquantique.com). In order to develop the research of quantum cryptography at the institute, close research cooperation has been established with industry: the "Latvijas Valsts radio un televīzijas centrs" LVRTC (www.lvrtc.lv), mobile operator LMT (www.lmt.lv), telecommunication company TET (www.tet.lv), and the Electronic Communications Office of Latvia (www.vases.lv). Currently, QKD technology has been tested in LVRTC and LMT fibre infrastructure. IMCS UL now implements the European Regional Development Fund project "Applications of quantum cryptography devices and software solutions in computational infrastructure framework in Latvia".

In the paper, authors describe the topics and tasks for deploying a cryptography digital ecosystem. Attention is devoted to descript methodologies and recommendation links in WEB that can contribute to establishing a quality of ecosystem.

2 Cryptography

Information plays a decisive and comprehensive role in our society. We need to obtain, transmit, and process information, but at the same time we also need to deny/hide access to information. Why access to data in the digital environment must be hidden:

- data has value and must be managed in order to use it for certain purposes;
- data has the privacy characteristics/rights of a person/organization;
- data availability ensures the security of the functionality of the digital envionment.

In a digital environment, access to information can be denied by locking information processing equipment (computers), for example, physically or using passwords. But data cryptography/data encryption (encryption) technologies are most widely used to control the availability of information. Encryption is a mathematical function that uses a secret value—a key that encodes data so that the information can only be read by a recipient who has access to that key. Encryption provides adequate protection against unauthorized or illegal data processing.

Information is encrypted and decrypted using one secret key (password).

The two main types of activity that require evaluating the use of encryption are da-ta storage and data transmission over physical data transmission networks.

Three cryptographic methods are used—using a secret-key or public key or a hash function solution.

The concept of cryptography, without delving into the nuances of modern cryptography, is essentially synonymous with the concept of coding, encryption, which ensures the transformation of information understandable to a person into its unintelligible content (unintelligible nonsense).

Modern cryptography can only be developed through the interaction of a large number of technical disciplines and it is also necessary to provide solutions related to data confidentiality, data integrity, and authorization [1].

3 Cryptography Versus Cybersecurity

The concepts of cryptography and cybersecurity are closely related and their difference lies in the nuances. Cybersecurity refers to the process of ensuring data security, while cryptography is one method of protecting sensitive information. However, cybersecrity and cryptography are two terms that should not be used interchangeably. The difference between the terms is visible in Tables 1 and 2.

Cryptography in everyday life includes several scenarios where the use of cryptography facilitates the provision of a secure service like email and file storage by using Pretty Good Privacy (PGP) freeware.

Five everyday uses of cryptography are as follows:

- encryption of company devices
- provision of e-mail communications
- protection of sensitive company data
- database encryption
- providing a WEB site.

Table 1 Cryptography versus cybersecurity concept

Cryptography	Cybersecurity
A system used to encrypt/decrypt data that cannot be understood by unauthorised users	The task refers to various measures taken by institutional organisations to detect and prevent malicious activities on networks or digital devices
A method of restricting access to information by unwanted persons. The encryption code and method is confidential	The method is not always effective in curbing cybercrime, as attackers can still bypass weak security systems
Technology used to improve cyber security	A set of activities where cryptography is only one option
Cryptography mitigates cybercrime using special technological solutions	Cybersecurity means maintaining specific procedures to ensure data security
Technology involves an aspect of personal knowledge, as the sender and recipient know each other's identity and are often familiar with the technological tool being used	Cyber security is not personal because its security policy is applied to a wide contingent of users

Table 2 Cryptography and cybersecurity origins

Purposes of Cryptography	Cybersecurity
Authentication Sender and receiver know each other's identity. They both know how to encrypt/decrypt the message and understand the intent of the message **Integrity** Data is securely stored and securely transmitted as no other user can access it. No other party can decipher the code and use or change the information **Confidentiality** An unauthorised user cannot understand the content of the data **Liability** in system actions or their origin. The perpetrator can then be held responsible for the offense **Non-repudiation** provides external evidence that the sender of the information is provided with proof of delivery, and the recipient with proof of the identity of the sender, so that later no one can deny that they have processed the information	• Network security • Data security • Application security • Security of mobile digital devices • Cloud computing security

4 Cryptography Digital Ecosystem Concept

"The digital ecosystem concept is designed to be similar to the existence of a natural ecosystem and has properties such as self-organisation, scalability and sustainability. Digital ecosystem deployment models are based on knowledge of natural ecosystems and the term is used in the computer and entertainment industries" [2].

In literature, you can find different ecosystem establishing models—on the basis of business relations and services [3] we find ecosystem models in higher education [4], but we deploy a Cryptography Digital Ecosystem model in connection with ICT cyber security and cryptography as a central technological solution for ICT security. The concept of Digital Ecosystem introduction started in 2000 at the World Economic Forum [5].

We follow the principles set forth by the World Economic Forum:

- "The digital renaissance and the global digital ecosystem". We draw parallels with the European Union (EU) Digital Europe Program 2021–2027 (DEP) [6].
- "Putting people at technology's heart". We draw parallels with ICT security priorities for everyone and everywhere, quantum computing and classical encryption col-lapse.
- "Partnership needs", we draw parallels with wide technology use and we indicate the cryptography technology belonging to General Purpose Technology line [7].
- "Sustainability", we draw parallels with DEP as an activity provider.

We define the digital ecosystem vertically (ecosystem status, in our case: EU level or national level) and horizontally—defining the boundaries that include technology (cryptography, in our case: ICT security and cryptography as a practical instrument for security needs).

5 Cryptography Digital Ecosystem Deployment Model Rationale

Let's introduce the facts that underpin the deployment of a digital ecosystem of quantum cryptography. Quantum cryptography is in EU development strategy with high priority.

5.1 QKD National Backbone Deployment

The Objective of the European Commission Digital Europe Programme (DEP) is the wide use of innovative technologies for igitalization of social life. DEP provides funding for projects in supercomputing, artificial intelligence, cybersecurity and advanced digital skills.

In this paper we will focus on cryptography, quantum cryptography and quantum communications as one of the methods to insure cyber security.

It can be considered that research in quantum computing with high priority in the EU started with the Quantum Flagship initiative launched in 2018. Research fields of this initiative cover fields of quantum computing, quantum simulation, quantum communication, quantum metrology and sensing core applications.

In 2019 European Union (EU) countries signed a declaration to explore together and deploy a quantum communication infrastructure (QCI) within the EuroQCI initiative.

In 2022 the European Commission announced three Calls as part of the initiative:

- Deploy advanced national QCI systems and networks;
- Create a European Industrial Ecosystem for Secure QCI technologies and systems;
- Coordinate the first deployment of national EuroQCI projects and prepare the large-scale QKD testing and certification infrastructure.

In this EuroQCI initiative, partners from Latvia (LVRTC, IMCS UL, TET, VASES) have presented Project LATQN and received a European Commission grant for the development of a national QKD (quantum Key Distribution) network back-bone as secure/restricted networking part and deployment of public QKD backbone part.

5.2 IPCEI on Next Generation Cloud Infrastructure and Services (IPCEI-CIS), Secure Priority

The European Commission's Strategic Forum for Important Projects of Common European Interest (IPCEI) initiative of DEP foresees contribution to economic growth. IPCEI projects are currently in the evaluation phase at the European Com-mission and at the second national phase.

IPCEI-CIS aims is to create a common cloud and edge infrastructure and its associated smart services for the future. IMCS UL's interest is directed at secure Cloud solutions, based on cryptography/quantum cryptography solutions.

A national consortium was created for the application of the project to the European Commission Call, which, during 2021, entered into an international partnership with OneFiber, Proemptor EC, and Engineering Ingegneria Informatica S.p.A. during the discussions organised by IPCEI. The current European Commission assessment of the joint document prepared by IPCEI can be found in IPCEI Annex 4, Chapeau Text on 31 March 2022.

The topic offered by Latvia within the framework of IPCEI was reduced to the pro-ject "Data protection, availability and processing solutions in next generation cloud infrastructure and secure communication technologies".

The objective of the project is to integrate quantum technologies into existing hardware and software infrastructure for CLOUD solutions. Also, the goal is to develop technologies for working with high entropy QKD and QRNG for their application in classic cryptographic engines, including integration into network equipment, thus allowing them to create a combined and fully functional solution that creates and maintains encrypted data transmission channels between remote objects with quantitatively synchronised keys, guaranteeing protection against information disclosure. IPCEI partners are also planned to be involved in the testing of the developed solutions. The project aims to provide the first industrial-scale solution for the integration of quantum devices in cloud and edge infrastructure and for shared access to QKD and entropy services.

The project's envisaged activities:

- integrate quantum technologies into existing hardware, software and infrastructure solutions; integrate post-quantum solutions into existing protocols; develop SDN OSI layer 2 business continuity; develop self-synchronising SDN channels
- integration of entropy as a service into existing communication networks and de-vices
- develop and integrate post-quantum secure solutions

Our project's subtopics is conceived as part of the IPCEI macro-project "Next Gen-eration Europe Cloud". The goal of the macro project is to develop unique software that will connect data centres across Europe and allow easy management of cloud resources and thus create a unified cloud.

Annex 4, Chapeau Text document includes subtopics:

- OneFiber and Macro project "X-Mesh: Application-Aware Edge-to-Cloud Stack" aims to define, validate, and document Quantum cryptographic technologies in the X-Mesh demonstrator, develop quantum cryptographic technologies that integrate into existing hardware infrastructure solutions, Application/Development of Aware Cloud Edge Infrastructure and usage patterns is developed on open usage plat-forms.

- The project “Software Defined Europe Wide Area Network” and the company “Pro-emptor EC” develop, build and integrate quantum cryptography technologies in SD EUWAN.
- The “Green Cloud Edge Federated Infrastructure” project of Engineering Ingegneria Informatica S.p.A. integrates container network interfaces with quantum cryptog-raphy technologies QKD, QRNG, Entropy-as-a-service.

5.3 IPCEI on Microelectronics Secure Priority

Objectives of microelectronics projects address to electronic equipment for information exchange with decision making (“Think”) properties.

Latvia mobile operator LMT (www.lmt.lv) participates in this initiative with the following ICT security development solutions:

- Highly secure quantum-based security systems based on Q-RNG;
- High-bandwidth, low-power consumption and secure next generation optical transceivers;
- Solution for 5G data connectivity devices.

IMCS UL as a national research partner takes part in those activities.

5.4 Is Cryptography a Widely Used Technology?

Technology is the application of research knowledge to practical aims.

Widely used technologies offer invention of a new product or process, application of the invention and multiply of this ideas [8, 9].

With the growing importance of cyber security, the widespread use of cryptography in technological platforms and security solutions is predicted.

5.5 General Purpose Technology (GPT) Line

Economists Richard Lipsey and Kenneth Carlaw introduced the term GPT [7].

The authors of the article have already analysed GPT in Latvia in 2008 [11].

If we analyse technology, it is important to understand how much it has an impact on society and how widely it is used. Now we value cryptography.

A general purpose technology always is associated with extensive use of innovative methods. Electricity and information technology (IT) are examples of GPT [10, 12].

Many authors recognise the internet as GPT. Blockchain can be recognised as the latest General Purpose Technology. See for example, the opinion in [13].

6 Cryptography Digital Ecosystem Deployment Framework

The digital ecosystem concept strongly relates to society's needs. Development of national level digital ecosystem frame limits are influenced by society, the political and economic system of the EU. The information we have provided, in our opinion, shows that the development of cryptography in Europe and Latvia has been widely evaluated and can be predicted to be widely used in the future.

The development of a large digital ecosystem depends on decisions made and fund-ing tenders announced. As a rule, the creation of a national digital system is based on several (many) participations in tenders, funding sources and several technological components. We believe that the national eco is created iteratively and is actually based on the evaluation of the EU society in advance, taking into account political and economic decisions.

7 Digital Cryptography Ecosystem Development Manual

The goal of the manual is to show, evaluate, summarise and predict the recommended uses of cryptography in the digital environment of Latvia. Also, the purpose of the manual is to recommend solutions/tasks and methodical approaches to the use of cryptography with the aim of improving IT security in Latvia. The content of the manual is divided into three sections with the following topics:

- Cryptography for us in everyday life: daily data privacy usage scenarios, customer survey;
- The readiness of the Latvian industry to implement cryptography, quantum cryptography platforms—from interviews of industry partners to solution projects;
- Technological aspects of cryptography: cryptography in communication networks and data processing protocols.

The design idea of the manual is borrowed from ENISA (https://www.enisa.europa.eu/). Since 2013, ENISA has published an annual document whose task is to provide the reader with summary news about the cyber security situation. In a similar way, the authors of the manual want to establish an overview of the use of cryptography as the main solution for ensuring cyber security for the Latvian public.

ENISA has accumulated rich experience in the preparation of such a document over many years and has formalised the document preparation process by preparing a methodology intended for action [15].

ENISA provides recommendations in the field of cyber security, but we, with the manual, focus more on one of the components of the implementation of cyber capability–cryptography.

The manual is prepared based on ENISA's recommendations and publications, however, the content presentation style is different. Our manual is more general and it is based on the above-mentioned approaches that are changing the content presentation style from "landscape" to "point of view" and "state of the art".

8 Cryptography Digital Ecosystem Development Objectives/Tasks/Topics Identified in the Manual

The list of ecosystem deployment strategy must cover a wide range of objectives/ tasks/topics:

1. Information access regulation, EU (international) level, society

- EU (international) information encryption regulatory principles: may/may not, must be decoded upon request
- Laws in Latvia: openness, protection, reuse of public information
- Information encryption and human rights
- Scenarios for everyday work in the digital environment when deciding to use cryp-tography, EU regulation reflection to society in Latvia
- eIDAS on electronic identification and trust services
  - Electronic signatures
  - Website authentication certificate (different from an electronic signature or seal certificate)
  - Electronic seals
  - Electronic time stamps
  - Registered secure electronic services
  - Registered secure electronic service certificate
- Passwords for data protection
- NIS directive about network and information systems security
- General Data Protection Regulation and data anonymisation.
  - Several anonymisation solutions are distinguished

    Deterministic pseudonymisation – the same pseudonym is always used for the same data;
    Case pseudonymisation in the document – using the same pseudonym for the same data only within the framework of the document;
    Completely random pseudonymisation – always using a different pseudonym for the same data.

  - Technology: nickname generator, Counter/Random number/Hash function/ HMAC/Encryption
- State registers, public data, geospatial data, databases, internet
- e-mail and encrypted attachments

- Secure web service usage scenarios
- Backups
- Cloud computing
- Regulation enforcement monitoring in ecosystem
  - National surveys
  - Analytical compilation of findings of international surveys

2. Industry readiness to implement cryptographic platforms

- QKD technology testing in real (LVRTC and LMT) fibre infrastructure
- EuroQCI tendering
- Establishing of the national QKD backbone

  Public sector
  Secure government sector

- IPCEI cloud secure applications, tendering
- Education and research projects

3. Technology development

- Study on the use of cryptographic techniques in Europe
  - ECRYPT, http://www.ecrypt.eu.org
  - ECRYPT II European Network of Excellence in Cryptology – Phase II, H2020
  - Crypto Service Gateway
  - Study on the use of cryptographic techniques in Europe, ENISA, 2012. Survey, 2014
- QKD networking (distance, key exchange rate)
  - Point-to-Point networks
  - Multiusers networks
- Quantum cryptography at the communication standard OSI levels
  - ADVA Layer 1 security solution
  - Layer 2 security
  - THALES Layer 2 and 3
  - Idquantique Layer common solution
- Priorities in quantum technology research in the EU FP7, H2020 and DEP DIGITAL programs, ENISA research settings (years 2016 and 2022)
- Cyber security (meaning including cryptography) in the curricula of Latvia lassifyies
- Term “Cryptography engineering”: cryptography technology aspect
- Standardisation in cryptography (ETSI, ISO, NIST)
  - EU Rolling Plan for ICT Standardisation 2020
  - Migration of cryptographic solutions to PQC, PQC and NIST
- PQC Maturity Assessment Model

- Cryptography strategy in the institution, risk management, good practice (crypto policy)
- Checklists
  - Software security requirements checklist
  - Checkpoint checklists for Microsoft, Azure Security Best Practices
  - The University of Toronto checklists for cryptography and information lassify-cation and protection
  - Checklists: SANS, Amazon (AWS), Microsoft (Azure), Google (GCP)
  - Security checklists for system requirements
  - Checklist for the cryptography designer at the institution
  - Checklist when hiring a cryptography specialist.

9 Conclusions

1. The authors promoted the national cryptography digital ecosystem deployment concept. In the view of the authors, this corresponds to DEP on QKD national backbone and international connectivity of these backbones.
2. Developed an example of national ecosystem readiness concepts.
3. Proposed a cryptography digital ecosystem deployment concept, that includes analysis of networking protocols, QKD networking, and QKD in OSI layer protocols.

Acknowledgements Publication vas supported from European Regional Development Fund project "Applications of quantum cryptography devices and software solutions in computational infrastructure framework in Latvia", Project ID number 1.1.1.1/20/A/106 (01.06.2021–30.11.2023).

References

1. Delaney J, Fireship LLC homepage, https://fireship.io/lessons/node-crypto-examples/. Accessed 05 Sept 2022
2. Digital ecosystem, Wikipedia, https://en.wikipedia.org/wiki/Digital_ecosystem#cite_note-9. Accessed 05 Sept 2022
3. Krohmer D, Naab M, Rost D, Trapp M (2022) A matter of definition: criteria for digital ecosystems, https://www.sciencedirect.com/science/article/pii/S2666954422000072. Accessed 05 Sept 2022
4. Wang Z, Zhang Q (2019) Higher-education ecosystem construction and innovative talents cultivating. Open J Soc Sci 7(3), https://www.scirp.org/journal/paperinformation.aspx?paperid=91072
5. Fiorina C (2000) The digital ecosystem, world resources institute conference: creating digital dividends. Seattle, Washington, October 16, 2000, http://www.hp.com/hpinfo/execteam/speeches/fiorina/ceo_worldres_00.html. Accessed 29 Aug 2022
6. The Digital Europe Programme, https://digital-strategy.ec.europa.eu/en/activities/digital-programme. Accessed 29 Aug 2022

7. Lipsey R, Carlaw KI, Bekhar CT (2005) Economic transformations: general purpose technologies and long-term economic growth. Oxford University Press. pp 131–218. ISBN 978-0-19-928564-8
8. Technology|Definition, examples, types, & facts, Britannica https://www.britannica.com›technology›technology. Accessed 29 Aug 2022
9. What is technology?—Definition & types, science courses / science fusion intro to science & technology: online textbook help, https://study.com/academy/lesson/what-is-technology-definition-types.html. Accessed 29 Aug 2022
10. General-purpose technology, Wikipedia, https://en.wikipedia.org/wiki/General-purpose_technology. Accessed 29 Aug 2022
11. Balodis R (2008) Rezga skaitlosanas izaicinajums latvijas zinatnei, Latvijas Zinatnu Akademijas Vestis 6, lpp 5–12
12. Jovanovic B, Rousseau PL (2005) General purpose technologies, handbook of economic growth, vol 1, Part B, chap. 18, pp 1181–1224, https://www.sciencedirect.com/science/article/abs/pii/S157406840501018X. Accessed 29 Aug 2008
13. Ozcan S, Unalan S (2022) Blockchain as a general-purpose technology: patentometric evidence of science, technologies, and actors. In: IEEE transactions on engineering management 69(3):792–809, https://ieeexplore.ieee.org/document/9166563. Accessed 08 Sept 2022
14. ENISA threat landscape (2021) April 2020 to mid-July 2021, https://www.enisa.europa.eu/publications/enisa-threat-landscape. Accessed 08 Sept 2022
15. ENISA cybersecurity threat landscape methodology (2022), https://www.enisa.europa.eu/publications/enisa-threat-landscape-methodology/@@download/fullReport. Accessed 08 Sept 2022

Exploring Out-of-Distribution in Image Classification for Neural Networks Via Concepts

Lars Holmberg

Abstract The currently dominating artificial intelligence and machine learning technology, neural networks, builds on inductive statistical learning processes. Being void of knowledge that can be used deductively, these systems cannot distinguish exemplars part of the target domain from those not part of it. This ability is critical when the aim is to build human trust in real-world settings and essential to avoid usage in domains wherein a system cannot be trusted. In the work presented here, we conduct two qualitative contextual user studies and one controlled experiment to uncover research paths and design openings for the sought distinction. Through our experiments, we find a need to refocus from average case metrics and benchmarking datasets towards systems that can be falsified. The work uncovers and lays bare the need to incorporate and internalise a domain ontology in the systems and/or present evidence for a decision in a fashion that allows a human to use our unique knowledge and reasoning capability. Additional material and code to reproduce our experiments can be found at https://github.com/k3larra/ood.

Keywords Trustworthy machine learning · Explainable AI · Neural networks · Concepts · Out-of-distribution

1 Introduction

Digitalisation influences all parts of society and central in this transformation resides artificial intelligence (AI) and machine learning (ML), technologies with roots in natural science and, as a consequence, a third-person objectivising stance [1]. Research in the area then inherits values that are concerned with average case metrics in the form of ground truth, optimising class probability, minimising bias in training data and mitigating consequences of data drift. This approach, useful when the target

L. Holmberg (✉)
Department of Computer Science and Media Technology, Malmö University, Malmö, Sweden
e-mail: lars.holmberg@mau.se

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_13

Fig. 1 Why is this image classified as a Tiger by a ResNet50 model? The exemplar classified is in the centre; it is flanked by two saliency maps on each side, maps that each accentuates some aspects of internal knowledge representations deemed as important for the classification by the XAI-method

domain is static and well defined, is not well suited to promote decisions[1] actionable in a non-static real-world context [2]. These ML/AI systems are information processing systems that due to their complexity become black boxes [3] that promote decisions without presenting reasons in a human-understandable form. Answering a *how-question*, by associating inputs to an ML-system with a promoted decision, compared to, answering a *why-question*, actionable in the real world, are two very different challenges. In later years, negative consequences related to this information processing approach, void of reasoning and understanding, have become increasingly apparent [4, 5].

This paper focuses on inductive statistical learning approaches used in the currently dominating ML technology, neural networks. This is a technology that can learn from raw input data in the form of images, sound or text [6]. In this ML/AI approach, knowledge priors embedded in the system originate primarily from the selection of labelled training data. Our focus is on classification tasks for images picturing mundane visual objects, a setting in which the strengths and weaknesses of the ML-system are less obscured compared to a more demanding domain. We thereby construct a setting in which a human with domain knowledge can be expected to assess reasons for a decision presented by the ML-system and therefore over time and usage build trust and knowledge concerning the system's strengths and weaknesses [7].

Our goal with this research is to go beyond average case metrics towards explanations for singular classifications and investigate if and in what way available explainable artificial intelligence (XAI) methods are useful when the goal is to understand and evaluate a promoted decision's validity. Central to this quest resides an ability to identify out-of-distribution exemplars. Building trust in an ML-system depends on the ability to answer: 'Not able to generalise for this exemplar', instead of, for all exemplars, present a class probability. By selecting a mundane domain and a few carefully selected images, we compare and discuss the limitations of several state-of-the-art pretrained neural networks and a number of XAI-methods (see Fig. 1). Our research focus is not on the technology as such, instead, it is to constructively uncover limitations related to the usage of neural networks and thereby discuss gen-

[1] The output from an ML-system in the form of classification, recommendation, prediction, proposed decisions or action.

erative directions that aim at building ML-systems that can produce, not only a label but, evidence for a decision actionable in a target context.

In the study, we find that a clearer distinction between labels promoted by the system and concepts, the human mental representation they refer to, is a generative way forward. Surrounding the neural network with a theory that can be used to identify o.o.d exemplars combined with an ontology for, the concept the label refers to, the subordinate concept is central for an ability to produce actionable explanations. Research directed towards concept learning and XAI-methods exposing subordinate concepts is a promising path forward, a path that can build on generalisable and well-defined concepts like basic shapes and patterns.

In the background section that follows, we present theories related to explanations, inductive statistical learning and concepts. This is followed by a section on the methodological approach leading to our study results. The article ends with a discussion section that contextualises our findings. We then end the article by concluding the results.

2 Background

Human understanding of natural phenomena builds on the uniformity principle which implies that instances of which we have no experience must resemble those we had experience from [8]. The uniformity principle cannot be justified from a scientific perspective and we, therefore, need to define non-inductive reasons to classify an explanation as scientific. To be scientifically valid, we need a theory that allows for deductive reasoning that consequently can be used to predict future events without relying on induction. This is then important for machine learning, based on induction, since without the ability to falsify the system and scientifically explain decisions, there is no demarcation between these systems and pseudoscience.

To define an explanation, we use Hilton's [9] definition that states 'the verb *to explain* is a three-place predicate: *Someone* explains *something* to *someone*'. A definition that focuses on explanations as a conversation between the explainer and the explainee. Additionally, the need for an explanation in a human context is often triggered by an event that is abnormal or unexpected from a personal point of view [10–12]. Research in explainable artificial intelligence (XAI) [13–18], on the other hand, is less concerned with *who* gives the explanation, to *whom* it is given or *why* it is needed [19] and has, in comparison, a more objectivising decontextualised stance to explanations.

In algorithmic ML, a theory is used to select an algorithm that can encompass the phenomena in question and, based on the assumptions and limitations of the algorithm selected, define the domain wherein predictions are valid. A neural network, instead of using a preselected algorithm, forms internal knowledge representations during training based on the tension between training data and labels.

In this work, we denote all human knowledge added to the ML-system as knowledge priors. For a neural network, knowledge priors are mainly added by network architecture, domain selection and input data selection. Neural networks then build internal knowledge representations from raw data (images, sound or text) and labels; the function used for prediction is then created in an inductive statistical manner by exposing the neural network to a large amount of training data [6]. If the training data and the labels are incalculable, the consequence is a non-transparent system [3].

The often assumed prerequisite for neural network training is that data and labels are independent and identically distributed (i.i.d), a presumption that is very challenging to fulfil in a real-world setting [20]. Exemplars consisting of data not part of the intended target domain are in our work denoted as out-of-distribution data (o.o.d.). A consequence of the inductive learning process is that o.o.d. data cannot be identified as external by a neural network, in line with the uniformity of nature principle. This since the network, cannot, without human help, identify the domain borders. It is then, in a classification problem, not possible to identify new classes; instead, the class probability for a promoted decision reflects some aspect of similarity with existing classes.

Concepts are central to human reasoning and essential for our ability to generalise. We use and need them to explain and make predictions about new objects and situations [21, 22]. Humans' beliefs, related to concepts and their properties, can be both false and incomplete and, additionally, contain both causal and descriptive factors [23]. A traditional approach used to search for a precise definition of concepts is to specify them as necessary and/or sufficient [22, 24]. This approach is used in machine learning both in experimental research [25] and to underpin representation learning [26]. To exemplify, the concept of 'elephant' can be described using necessary subordinate concepts shared with many animals, for example, 'four legs', 'eyes' and sufficient subordinate concepts, for example, 'elephant tusk'. Sufficient and necessary subordinate concepts are those that on their own can be used in classification, for example, 'elephant trunk' ('elephant tusks' are sufficient for classification but not necessary since not all female Indian elephants have tusks). Spurious correlations are relationships between the proposed decision that are prone to change when a system is deployed in a real-world context [27]. For example, if all elephants are pictured close to 'watering holes', this concept can wrongly be seen by the ML-model as a necessary concept. In one part of our study, we investigate the usefulness of spurious correlations and sufficient and necessary concepts since human insights here can help to identify limitations in the training data in relation to the deployment domain. In later research, classification of concepts by the use of necessary and sufficient is replaced by prototype theories implying a somewhat looser definition to avoid contradictions and instead define a concept as central tendencies of the phenomena in question [22]. In this work, we use the notion of labels synonymous with referents, prediction, promoted decisions and classifications, and we differentiate them from concepts that we define as the human mental representation referred to by a label. In this article, the writing context clarifies if we refer to a label or a concept, but, when we find it important to make a distinction clear, we surround the label with double quotes and the concept with single quotes.

In recent years, there has been a surge in XAI-methods [16, 28], but there are few user studies evaluating these methods in a real-world context [29]. For our study, we have selected both XAI-methods that are ML-model independent, like Occlusion [30] and those that focus on internal knowledge representations in neural network layers closer to the promoted decision, two Gradient-weighted Class Activation Mapping (Grad-CAM) [31] methods and a SHapley Additive exPlanations (SHAP) [32] method, as well as one method, integrated gradients (IG) [33], that weighs activations over all layers in a neural network. XAI-methods, in their original papers, are often presented with best-case performance examples and their usefulness can even be questioned [34]. We are then interested in using the methods to evaluate their usefulness when the goal is to identify o.o.d exemplars. For the studies we selected, on ImageNet-1K [35], pretrained models[2] both for our initial qualitative study and for the controlled experiment in which we compare eight models.

3 Methodology

We set up a targeted case study [36] to explore some alternative approaches that ML/AI systems can use to communicate reasons for decisions with humans. Our goal is to put the technology and not humans on the test bench and ask questions related to which research paths are feasible to make this technology useful if the goal is human understanding related to singular classifications. One of our studies investigated if the notion of sufficient concepts, necessary concepts and spurious correlations can be useful as an analytical tool to categorise areas in images accentuated by the XAI-methods and thereby underpin understanding. In the second study, we investigated if we can rely on a human intuitive understanding of the XAI-methods explainability capabilities. In both cases, we selected images picturing non-abstract concepts clearly visible in the images (animals and headgear). The studies mixed, for the ML-model and humans, harder-to-identify objects with easier ones and a few images pictured more than one object (see the accompanying website). We selected images picturing objects predicted with class probability around 50% so the probability would not take precedence over the XAI-methods. The study participants were not told which classes the system could recognise, they only got information related to the domain they could expect, animals or headgear. By this approach, we aimed towards exploring the usefulness of the XAI-methods combined with classification metrics to identify o.o.d. exemplars in a mundane domain.

Our third study aims towards a more objective study comparing the behaviour of pretrained models in relation to a specific concept: ‘sorrel’. In this controlled experiment, we investigate if pretrained ML-models are aligned in their predictions and if areas in the image are, by the model agnostic XAI-method Occlusion, denoted as equally important between the models. We selected ‘sorrel’ as a typical human non-abstract concept that additionally is represented as a class in ImageNet-1K. By doing this, we can concretise a discussion on what we can expect, and not expect, from classification systems deployed in a real-world context.

[2] https://pytorch.org/vision/stable/models.html.

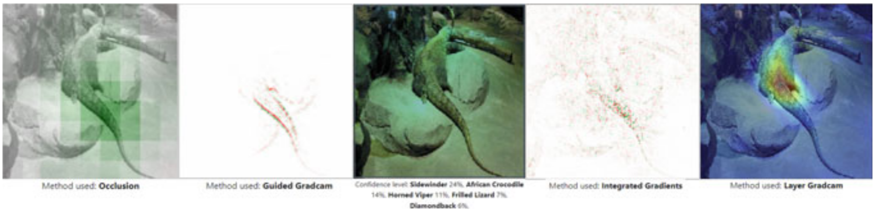


Fig. 2 Examples of the diverse visual languages used in XAI-methods. The original image and predictions are placed in the middle, flanked by different XAI-methods

We are aware that ImageNet-1K is part of a competition to optimise average case metrics (ILSVRC [35]); still, we find it useful to discuss singular classifications for these ML-models since the models are a blueprint for models deployed in real-world settings. By focusing on these, from an average case perspective, state-of-the-art ML-models, we aim at illuminating deficiencies that follow an objectivising stance. Our focus is based on a human in command relation to AI/ML and, consequently, a focus on subjective understanding [37]. XAI, to be useful, needs as a minimum to function as a trustworthy tool for our selected group of younger persons with IT-related education, if we are to expect them to be useful for other groups. One important ingredient, that we focus on via o.o.d, is a system where it is humanly possible to evaluate if a classification makes sense and can earn a user's trust.

By using well-known datasets, a coherent XAI-API,[3] pretrained models and published code, our aim is to make our study reproducible, and thereby, our results generative for similar studies. For the two user studies, we created a website that used a pretrained ResNet50 network, a model with a reasonably low error rate (@1 is 76.1% and @5 92.9% on the ILSVRC [35] challenge). Each study consisted of web pages the participant could navigate back and forth between. On each page, one image pictured an object and a prediction together with the possibility to choose an XAI-method to investigate. A form to collect more structured answers was placed under the images accompanied by free text fields. After the study, the participants were asked to summarise their understanding of the ResNet50-models overall capabilities related to the domain in question. By keeping many parameters constant, we aimed towards a controlled experiment [38] and semi-structured interviews [39]. For the user studies, we selected, by convenience sampling, ten participants from IT-related education at the bachelor level. The age span was 20–40 years, and 40% of the participants were women. Except for collecting the form data from the website, we interviewed six of the participants in half-hour sessions. The studies were discontinued when we found that they saturated.

The XAI-methods selected, in their original implementations, use a diverse and to some extent incompatible graphical language. For example, Grad-CAM images are often visually appealing compared to gradient-based methods (see Fig. 2). Differences lay partly in the colour schemes used and if the visual explanations are overlaid

[3] https://captum.ai/.

on the original image. Grad-CAM uses, in the original implementation, colour gradients from blue to red to indicate areas that increasingly influence the prediction. Other methods use red colour to indicate areas that influence the attribution negatively [25]. These diverse visual explanation languages are then not comparable and each of them needs to be explained to be understood. We discussed, in our research team, what a negative attribution implies and came to the conclusion that it is not intuitive and its usefulness for our experiment can be questioned. The main reasoning is that negative attribution for one class implies that it is more positive for an unknown amount of one to many of the other 999 classes, thereby opening up for complex contrastive speculations. Instead of adding to the complexity more than needed, we decided to focus on identifying subordinate concepts that can be associated with the promoted decision, whether the prediction is perceived as correct or incorrect by the user. Based on the reasoning above, we decided to use a coherent graphical language for the XAI-methods and thereby increase comparability between the methods. In the user studies, we presented two images, the original image with predictions and another image with the selected XAI-method overlaid on the original image in black and white and slightly opaque. We then used a bright green colour that contrasts with the greyish background, and we let the opacity of the green colour indicate how important a part of an image is for the promoted decision (see Figs. 1 and 3). We also decided for Grad-CAM and Occlusion to use 7×7 squares to make it easier to reason about subordinate concepts.

For the third study, we only used Occlusion [30] since it is model agnostic and therefore visualises the importance of each square in a fashion that makes it possible to compare the ML-models more objectively. The different architectures of ML-models make it hard to use model-dependent XAI-methods to compare models since they depend on how the internal knowledge representation is structured. The method Occlusion can then be seen as more objective since it mechanically measures, somewhat simplified, how important individual squares in a grid, overlaid on an image, are for a promoted decision, and we can therefore compare the focus for the different models.

4 Result

In the first study presented below, more complex images picturing primarily animals, whilst the second study has a more narrow focus on headgear. The third study compares the agreement between eight pretrained models for predictions in relation to the concept 'sorrel'.

4.1 Out-of-Distribution Using the Notion of Necessary and Sufficient

One of our three studies investigated whether it is useful, for a human, to categorise visible subordinate concepts in images picturing animals as necessary, sufficient and the usefulness of the notion of spurious correlations. By using these notions we, for this study, hypothesise that they can make it easier to identify o.o.d exemplars. Concepts discussed here were then subordinate concepts to animal classifications, as, for example, ‘watering hole’, ‘beak’ or ‘feather’.

The methods Layer Grad-CAM and Guided Grad-CAM were in the study deemed as the ones that resemble a human approach closest and thus were then denoted as most useful. For example, a somewhat non-sharp image picturing a type of lizard ‘komodo dragon’ that is among the ImageNet-1K classes was by the ML-model erroneously classified as a type of snake with 24% class probability (see central image in Fig. 2). In this case, the XAI-methods were useful since drew attention to the form of the tail as a possible reason, that for a human, can result in some generalisable knowledge related to the ML-models behaviour. Another example: since ‘horse’ is not one of the classes in ImageNet-1K, it is not possible to classify for the ML-model. This was disclosed by two users that used the XAI-methods to point towards reasons why ‘horse’ was erroneously classified as the dog bread “great Dane” and thus drew the conclusion that ‘horse’ must be an o.o.d class (see Fig. 4 and below for discussions concerning horse and sorrel).

The study participants judged from their personal knowledge that the ML-model predicted correctly for 60% of the images. In total, 64 images were assessed by the participants. The perceived usefulness of the XAI-methods for the animal study is presented in Table 1. The fact that the study participants did not know which exemplars that were o.o.d was found to be confusion. For example, classifying a children’s pool with three ducks as a drake is from a human perspective a questionable focus but technically understandable, given that other classes related to ducks, pools or ponds are not among the available classes. The use of necessary concepts, sufficient concepts and spurious correlations can probably be useful but our results indicate that the study participants need a theoretical base to make use of these notions. For the majority of the users, this approach was not seen as a useful path to better understanding reasons for a promoted decision. Only one participant was cautiously positive, but generally, the usage of this categorisation of concepts was seen as puzzling or even uncomfortable since the notions are subjective, sometimes contradicting and open for discussion.

4.2 Out-of-Distribution with Headgear

The other part of the study focused on a more narrow domain ‘headgear’ and the seven directly related classes in ImageNet-1K: ‘sombrero’, ‘cowboy hat’, ‘bathing cap’,

Fig. 3 Part of the headgear study, visualising areas in images accentuated as important for a promoted decision. The original image and predictions are placed in the middle, flanked by different XAI-methods

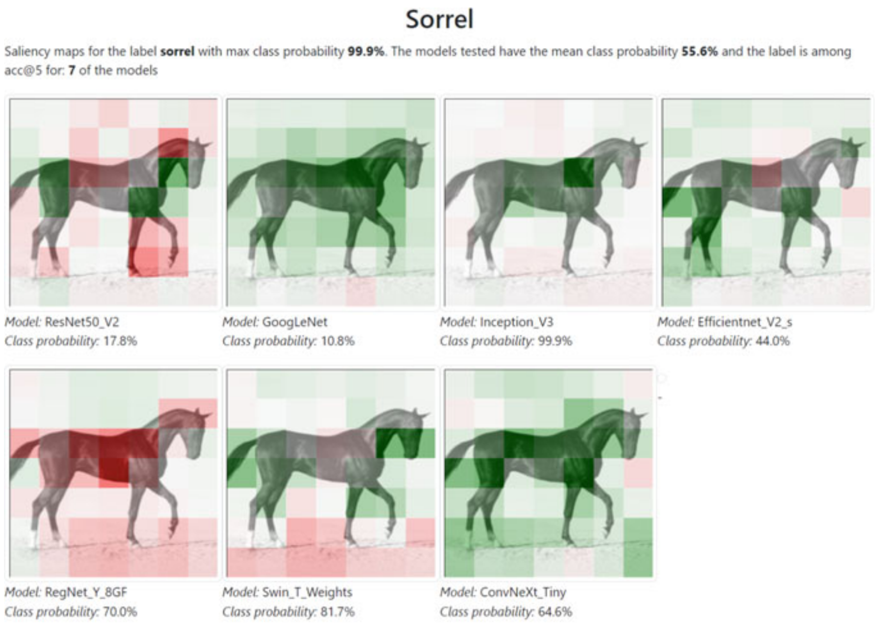
Table 1 The participant's subjective and relative assessment of the usefulness of XAI-methods

	Animal study			Headgear study		
XAI-method	TP (%)+ FP (%)	TP (%)	FP (%)	TP+FP (%)	TP (%)	FP (%)
Occlusion	19	18	19	29	29	31
Guided grad-CAM	35	31	42	27	27	27
Integrated gradients	7	5	12	5	5	6
Layer Grad-CAM	38	45	27	36	38	32
Gradient SHAP	n/a	n/a	n/a	3	2	4

TP (True Positives) denotes usefulness of XAI-methods for predictions perceived as correct. FP (False Positives) denotes the usefulness of methods for predictions perceived as erroneous (The participant could choose multiple XAI-methods for each proposed prediction)

'crash helmet', 'bonnet', 'shower cap' and 'football helmet'. In this study, we did not use any concept theory and instead relied on the participant's intuitive understanding of headgear-related concepts. This closer adheres to prototype theories and concepts defined as central features of the phenomena in question [22].

The study participants judged that the ML-model predicted correctly in a little bit more than half of the cases (55%). We found only a slight difference between the perceived usefulness of XAI-methods for the classifications that were judged as correct and those judged as incorrect. The perceived usefulness of the XAI-methods for the headgear study is presented in Table 1. In total, 115 images were assessed by the participants. Examples of the perceived usefulness are, for example, related to an image picturing a 'stormtrooper helmet' from the Star Wars movies that were classified as a "crash helmet". This was for some users seen as correct since it resembles a 'crash helmet' by the areas accentuated, especially by, Guided Grad-CAM. The similarity in functionality was discussed by one user in that a 'stormtrooper helmet' most probable has a protective function but that particular participant also concluded that the superordinate concept 'helmet' (not among ImageNet-1K classes) would fit better. In this test, the participant became aware of missing classes, for example, that a 'top hat' was classified as a "cowboy hat" (51% class probability), opened up for these types of speculations. Other more general opinions were that the ML-model seemed to be good at fabric for example 'wool' but also biased towards water-related



ResNet50_V1

Prediction	%
Great Dane	41.4
whippet	15.8
Doberman	13.9
Mexican hairless	6.0
Staffordshire bullte...	3.2

ResNet50_V2

Prediction	%
sorrel	17.8
bell pepper	3.7
cowboy hat	0.7
Great Dane	0.7
Weimaraner	0.6

GoogLeNet

Prediction	%
Great Dane	12.2
sorrel	10.8
Rhodesian ridgeback	8.7
black-and-tan coonho...	6.7
Doberman	6.0

Inception_V3

Prediction	%
sorrel	99.9
horse cart	0.0
cowboy hat	0.0
book jacket	0.0
worm fence	0.0

Efficientnet_V2_s

Prediction	%
sorrel	44.0
worm fence	2.8
cowboy hat	1.5
muzzle	0.8
cowboy boot	0.5

RegNet_Y_8GF

Prediction	%
sorrel	70.0
horse cart	0.5
Doberman	0.2
hartebeest	0.2
ox	0.2

Swin_T_Weights

Prediction	%
sorrel	81.7
seashore	0.2
Saluki	0.2
worm fence	0.1
red-breasted mergans...	0.1

ConvNeXt_Tiny

Prediction	%
sorrel	64.6
Great Dane	2.4
cowboy hat	0.8
black-and-tan coonho...	0.6
Doberman	0.6

Fig. 4 Five Occlusion saliency maps images at the top indicates with green colour shade the positive impact that specific square has for the prediction. Red colour indicates that the square has negative impact on the prediction. The colours are normalised and cannot be compared between the images. Under the images, top 5 class probabilities for the eight models are presented

headgear ('bathing cap', 'bonnet', 'shower cap') especially when these were worn by people (spurious correlation).

4.3 Comparing Models

In this part, we make a comparative analyse of predictions and XAI-explanations related to an image picturing a black horse, an image that also was used in our first study. Using eight pretrained models and the model agnostic XAI-method Occlusion, we compare and discuss predictions from an o.o.d perspective. The predictions can be seen in Fig. 4. The only model that does not predict sorrel as one of the top 5 classes (@5) is the model we used in Study 1 (ResNet50 V1). According to the dictionary Merriam-Webster.[4] sorrel has two definitions either, it is a light bright chestnut-coloured horse or a plant with sour juice, typically common sorrel (Rumex acetosa). These two different concepts, a type of horse and a group of plants carries a wealth of knowledge that, we as humans, connect to our real-world knowledge. For example, if a person is knowledgeable about horses, they will connect this horse colour with the horse-breed quarter-horse and that this horse in Europe commonly is denoted chestnut. And, of course, the cultural global-north discourse connected to these labels and concepts can be taken into account depending on whom the classification should be useful for. The type of plants denoted as sorrels similarly follows a wealth of causal and descriptive factors that also can be contextualised. In WordNet [40] (the taxonomy ImageNet is based on) semantic relations to sorrel, additionally, adds a definition of sorrel as an adjective for a brownish-orange colour. In this work, we lift out **sorrel** as an example of a concept that for humans, and for the ML-models tested, is incomplete, contextual and contains, both causal and descriptive factors [23].

Sorrel is among @5 for seven of the models tested and @1 for six of them (see Fig. 4). For ResNet50, using W1 weights as in our first study, "sorrel" is not among the five top predictions and the model instead generalises towards dog breeds. The other models generalise towards sorrel for the black horse, and by studying the class probabilities, we can hypothesise on training data distribution. The saliency maps show that the ML-models focus on different parts in the image to classify the horse as a "sorrel". Therefore, our study participants in the first study concluded that horse was not part of the target domain, even if the horse subordinate concept sorrel is. If we would have used any of the other models, the participants would likely, if they did not know what a sorrel horse is, incorporate that into their knowledge of subordinate concepts connected to 'horse'.

The plant-related concept connected to the label sorrel is not picked up by the models and can be concluded to be an o.o.d. for the training data. If the eight models are exposed to an image of the plant species common sorrel (see Fig. 5), it is classified as ear (@1) by all models with a mean class probability of 55.7%. The label "ear" is part of the training data using the plant-related definition 'the fruiting spike of a cereal (such as wheat or corn) including both the seeds and protective structures'.[5] which is misleading since a common sorrel belongs to a different natural group than plants with 'ears'. The more common usage of the concept 'ear', as a hearing organ for vertebrates, is not represented in the training data for the ImageNet class "ear".

---

[4] https://www.merriam-webster.com/dictionary/sorrel Accessed 5 Sep. 2022.

[5] https://www.merriam-webster.com/dictionary/ear Accessed 7 Sep. 2022.

ResNet50_V1

Prediction	%
ear	60.4
corn	22.1
hay	2.4
coral fungus	2.0
harvester	1.4

ResNet50_V2

Prediction	%
ear	57.7
corn	28.7
hay	2.5
groom	0.2
limpkin	0.1

GoogLeNet

Prediction	%
ear	32.8
corn	15.0
rapeseed	2.5
prairie chicken	2.4
bittern	2.3

Inception_V3

Prediction	%
ear	53.5
corn	46.4
cardoon	0.0
king crab	0.0
coral reef	0.0

Efficientnet_V2_s

Prediction	%
ear	49.2
corn	22.6
hay	0.3
yellow lady's slippe...	0.2
lakeside	0.2

RegNet_Y_8GF

Prediction	%
ear	68.2
corn	4.0
hay	0.6
birdhouse	0.3
harvester	0.3

Swin_T_Weights

Prediction	%
ear	59.2
corn	20.3
cardoon	2.3
hay	1.8
long-horned beetle	1.0

ConvNeXt_Tiny

Prediction	%
ear	64.5
hay	2.6
corn	2.2
bittern	0.9
coral reef	0.6

Fig. 5 We exposed the ML-models for an image of a common sorrel with the above-presented result. We can from the class probabilities draw the conclusion that the plant corn (that has ears) is part of the training data since the models are biased in that direction

4.4 Concluding Remarks on the Studies

Our studies show that the usefulness of model predictions and XAI-method depends on a human ability to compare the evidence exposed with the reality as perceived subjectively by humans. In our setting there is not much to learn about the reality for humans, usefulness resides in a deepened understanding of the ML-model's behaviour related to the domain exemplars belong to. These insights can then be used to improve the ML-model and/or delimit its usage domain. It is also worth noting that we in our two first studies measure perceived and relative usefulness in relation to the other XAI-methods in the study. The only conclusion we can draw is that, for these mundane images, some XAI-methods are perceived as better than others, and it is hard to not distinguish what is objectively better from those perceived as better due to confirmation bias. The lack of a taxonomy that relates concepts to each other also becomes evident in the comparative study. Here, a closer integration with WordNet [40] or a similar service could add further insights useful for a human when predictions are analysed. Our main take-off from our studies is that focus moves to the context in which the system is to be used and consequently what the question was.

5 Discussion

In this section, we use our results to discuss challenges in relation to actionable explanations for singular image classifications. By adhering to falsification as a fruitful research avenue, we aim to lay bare future research paths.

For the images used and the participants we selected, we found that the two Grad-CAM methods correlate better with human understanding than the other XAI-methods in the study. The gradient methods were in our study deemed less useful than the other XAI-methods. The methods and class probabilities taken together made it possible for study participants to speculate on training data distribution, but it is hard to draw concrete conclusions. We did, additionally, not find that the notion of necessary and sufficient concepts and spurious correlations added any clarity. In the light of our third study and Occlusion saliency maps produced by the different ML-models, we draw the conclusion that the internal knowledge representations for the ML-models are not aligned between the models and that we therefore cannot expect more clarity for end users if we use other models.

It therefore becomes evident, even through a limited study like this, that these systems lack a theory that can be used deductively. A theory covering, to the labels, subordinate concepts combined with a theory that couples superordinate concepts to the target context. Explanations in the decontextualised setting we constructed for our studies can only give insights into strengths and limitation for the ML-model in relation to the training data and not explain something to someone that is directly actionable in a real-world context [9].

The flat output hierarchy of labels from the neural networks we used combined with assuming independent and identically distributed training data adds to the complexity we discuss. For example, in the studies involving the concept ‘sorrel’, we elucidate that for humans, concepts are often incomplete, false and contain both descriptive and casual factors [22]. In this study, we show that this complexity is lost when the concept is reduced to 2d images and strings of characters. Additionally, as a consequence, there is no logical reason that internal knowledge representations in latent spaces and variables are aligned with, to a predicted label, subordinate concepts. Consequently, we cannot expect these systems to create actionable explanations, instead what we can expect, sometimes, is that they can help us draw attention to correlations hard for humans to uncover with our senses. An obvious problem is then that the correlations partly will overlap with human understanding of the domain and that the internal knowledge representations in the neural network build an alternative black-boxed ontology.

For a scientific field like biological classification, superordinate classes can be used to delimit the target domain to avoid that a user expose the system to o.o.d exemplars. When it becomes ethically trickier, like classifying within the people sub-tree of ImageNet [41], the target domain, the training data and the labels need to be synchronised. If these systems are out of sync with the context in, for example, a transfer learning setting, they transfer internal knowledge representations that hide norms and values from one context to another. This can from our study be exemplified with our 'sorrel' example that both populate our ontological understanding of the concept and, to some extent, at the expense of other concepts. A more worrying example is the work by Bender et al. [42] coining the concept of 'stochastic parrots' for large language models.

Research in the concept direction is increasingly attracting attention. For example, instead of end-to-end training *concept bottleneck models* [43] are trained on subordinate concepts that can be used in explanations. A related approach is concept activation vectors (CAV) where user-defined subordinate concepts are defined by representative exemplars and used to measure the label's sensitivity for a subordinate concept [44]. Another approach uses faultlines to identify (using Grad-CAM) and remove or add, to the label, subordinate concepts and thereby underpin contrastive explanations [45].

The bulk of the work mentioned above implies learning subordinate concepts that in themselves are complex and therefore to a large extent move the challenge concerning concepts so they, in essence, keep the challenge. Examples here are evaluating health data [46] or classifying skin lesion [47] using complex subordinate concepts. Other examples, related to our study, involve detecting subordinate concept and combining them with deductive rules, for example, detecting spatial relations between face parts to identify faces [48] or using self-organised decision trees to classify images based on subordinate concepts like bed, sea or tree [49].

The research path we believe answers to recent research challenges, and that our findings point towards, is learning basic concepts that can be defined objectively and are widely generalisable, for example, basic shapes, colours and patterns. By using these concepts as building blocks for explanations, more complex explanations can be built. This since a combination of shapes, colours and patterns can be used to build relatively complex explanations grounded in concepts that can be agreed on. This leaves interpretation, reasoning and understanding implications to humans.

6 Conclusion

In this paper, we analyse the consequences of the inductive statistical learning process that underpin knowledge creation in neural networks. The learning process used implies that the systems cannot distinguish exemplars that are part of an intended knowledge domain from exemplars that are out-of-distribution. In our study, we focus on image classification and local XAI-methods and show that reasons for a decision have to be interpreted by a human with domain knowledge to be explainable. We also find that XAI-methods used in our study produce vague and incoherent reasons for a decision, reasons that additionally are open to different interpretations. When we analyse the incoherence, using the notion of concepts, we find that we need a better alignment between internal knowledge representations in the neural networks and, to the presented label, subordinate concepts.

For future work, we suggest, using basic and definable concepts as building blocks for AI/ML-produced explanations. Elaborate, colourful, precise and contextually relevant explanations, that humans can produce, can then be traded for trustworthy explanations.

Acknowledgements This work was partially financed by the Knowledge Foundation through the Internet of things and people research profile. I am also in debt to colleagues at Malmö University for their invaluable and insightful comments that substantially improved the work. Images used are either copyright free or used by permission.

References

1. Grimm SR (2016) How understanding people differs from understanding the natural world. Nous-Supplement: Philos Issues 26(1):209–225
2. Chollet F (2019) On the measure of intelligence, p 64. ArXiv preprint arXiv:1911.01547
3. Lipton ZC (2016) The mythos of model interpretability. Commun ACM 61(10):35–43
4. Hutchinson B, Mitchell M (2019) 50 years of test (Un)fairness: lessons for machine learning. In: FAT* 2019—Proceedings of the 2019 conference on fairness, accountability, and transparency, pp 49–58. Association for Computing Machinery, Inc
5. Couldry N, Mejias UA (2019) Data colonialism: rethinking big data's relation to the contemporary subject. Telev New Media 20(4):336–349
6. Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
7. Holmberg L (2021) Human in command machine learning. No. 16 in Studies in Computer Science
8. Henderson L (2020) The problem of induction. In: The Stanford encyclopedia of philosophy. Metaphysics Research Lab, Stanford University
9. Hilton DJ (1990) Conversational processes and causal explanation. Psychol Bull 107(1):65–81
10. Hilton DJ, Slugoski BR (1986) Knowledge-based causal attribution: the abnormal conditions focus model. Psychol Rev 93(1):75–88
11. Hesslow G (1988) The problem of causal selection. Contemporary science and natural explanation: commonsense conceptions of causality, pp 11–32. https://www.hesslow.com/GHNew/philosophy/Problemselection.htm
12. Hilton DJ (1996) Mental models and causal explanation: judgements of probable cause and explanatory relevance. Thinking Reasoning 2(4):273–308
13. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R, Herrera F (2020) Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion 58:82–115
14. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D (2018) A survey of methods for explaining black box models. ACM Comput Surv 51(5):42
15. Biran O, Cotton C (2017) Explanation and justification in machine learning: a survey. IJCAI workshop on explainable AI (XAI). 8:8–14
16. Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L (2019) Explaining explanations: an overview of interpretability of machine learning. In: Proceedings—2018 IEEE 5th international conference on data science and advanced analytics, DSAA 2018. IEEE, pp 80–89
17. Adadi A, Berrada M (2018) Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access 6:52138–52160
18. Hoffman RR, Clancey WJ, Mueller ST (2020) Explaining AI as an exploratory process: the peircean abduction model. ArXiv preprint http://arxiv.org/abs/2009.14795

19. Miller T (2019) Explanation in artificial intelligence: insights from the social sciences. Artificial Intelligence 267:1–38
20. Schölkopf B (2019) Causality for machine learning (2019). http://arxiv.org/abs/1911.10500
21. Margolis E, Laurence S (2021) Concepts. In: The Stanford encyclopedia of philosophy. Metaphysics Research Lab, Stanford University
22. Murphy G (2004) The big book of concepts. MIT Press
23. Genone J, Lombrozo T (2012) Concept possession, experimental semantics, and hybrid theories of reference. Philos Psychol 25(5):717–742
24. Brennan A (2017) Necessary and sufficient conditions. In: Zalta EN (ed) The stanford encyclopedia of philosophy. Metaphysics Research Lab, Stanford University, Summer 2017
25. Wang Z, Mardziel P, Datta A, Fredrikson M (2020) Interpreting interpretations: organizing attribution methods by criteria. In: IEEE computer society conference on computer vision and pattern recognition workshops, vol 2020-June, pp 48–55
26. Wang Y, Jordan MI (2021) Desiderata for representation learning: a causal perspective
27. Gulrajani I, Lopez-Paz D (2020) In search of lost domain generalization. http://arxiv.org/abs/2007.01434
28. Samek W, Montavon G, Lapuschkin S, Anders CJ, Müller KR (2021) Explaining deep neural networks and beyond: a review of methods and applications. Proc IEEE 109(3):247–278
29. Tjoa E, Guan C (2019) A survey on explainable artificial intelligence (XAI): towards medical XAI. In: IEEE transactions on neural networks and learning systems. https://arxiv.org/abs/1907.07374
30. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. Tech rep
31. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-CAM: visual explanations from deep networks via gradient-based localization. Int J Comput Vis 128(2):336–359
32. Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. In: Advances in neural information processing systems, vol 3, MIT Press, pp 4766–4775
33. Sundararajan M, Taly A, Yan Q (2017) Axiomatic attribution for deep networks. In: 34th International conference on machine learning, ICML 2017, vol 7, pp 5109–5118
34. Adebayo J, Gilmer J, Muelly M, Goodfellow I, Hardt M, Kim B (2018) Sanity checks for saliency maps. In: Advances in neural information processing systems, vol 2018-Decem, pp 9505–9515. https://goo.gl/hBmhDt
35. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) ImageNet: A large-scale hierarchical image database. IEEE conference on computer vision and pattern recognition. IEEE, Miami, pp 248–255
36. Seawnght J, Gerring J (2008) Case selection techniques in case study research: a menu of qualitative and quantitative options. Polit Res Q 61(2):294–308
37. Holmberg L (2021) Human in command machine learning. http://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-42576
38. Ko AJ, LaToza TD, Burnett MM (2013) A practical guide to controlled experiments of software engineering tools with human participants. Empirical Softw Eng 20(1):110–141
39. Myers MD, Newman M (2007) The qualitative interview in IS research: examining the craft. Inf Organ 17(1):2–26
40. Miller GA (1998) WordNet: an electronic lexical database. MIT Press
41. Yang K, Qinami K, Fei-Fei L, Deng J, Russakovsky O (2020) Towards fairer datasets: filtering and balancing the distribution of the people subtree in the ImageNet hierarchy. In: FAT* 2020—Proceedings of the 2020 conference on fairness, accountability, and transparency, pp 547–558
42. Bender EM, Gebru T, McMillan-Major A, Shmitchell S (2021) On the dangers of stochastic parrots: can language models be too big? In: FAccT 2021—proceedings of the 2021 ACM conference on fairness, accountability, and transparency, pp 610–623
43. Koh PW, Nguyen T, Tang YS, Mussmann S, Pierson E, Kim B, Liang P (2020) Concept Bottleneck models. In: International conference on machine learning (2020). https://arxiv.org/abs/2007.04612

44. Kim B, Wattenberg M, Gilmer J, Cai C, Wexler J, Viegas F, Sayres R (2018) Interpretability Beyond feature attribution: quantitative testing with concept activation vectors (TCAV). Tech, Rep
45. Akula A, Wang S, Zhu SC (2020) CoCoX: generating conceptual and counterfactual explanations via fault-lines. In: Proceedings of the AAAI conference on artificial intelligence 34(03):2594–2601
46. Mincu D, Loreaux E, Hou S, Baur S, Protsyuk I, Seneviratne M, Mottram A, Tomasev N, Karthikesalingam A, Schrouff J (2021) Concept-based model explanations for electronic health records. In: ACM CHIL 2021—proceedings of the 2021 ACM conference on health, inference, and learning, pp 36–46
47. Lucieri A, Bajwa MN, Alexander Braun S, Malik MI, Dengel A, Ahmed S (2020) On Interpretability of deep learning based skin lesion classifiers using concept activation vectors. In: 2020 International joint conference on neural networks (IJCNN), pp 1–10
48. Rabold J, Schwalbe G, Schmid U (2020) Expressive explanations of DNNs by combining concept analysis with ILP. In: German conference on artificial intelligence 12325 LNAI, pp 148–162
49. Elshawi R, Sherif Y, Sakr S (2021) Towards automated concept-based decision tree explanations for CNNs. In: Advances in database technology—EDBT, vol 2021-March, pp 379–384

Robust GNSS/Visual/Inertial Odometry with Outlier Exclusion and Sensor's Failure Handling

Bihui Zhang, Xue Wan, and Leizheng Shu

Abstract Improving the robustness of multisensor fusion state estimation is critical for applying these new techniques into practical applications. To this end, we propose outlier detection and exclusion methods for the visual data and GNSS data in a tightly coupled, sliding window optimization-based GNSS/visual/inertial odometry. To handle the complete failure of the visual data, we also propose a visual termination and keyframe fast recovery strategy. Real-world tests with multipath effect of GNSS signals and severe visual interferes are performed. Experimental results show the effectiveness of our methods in improving the robustness of the odometry, and increase of computation time due to the method are also analyzed. Attached video of the tests is available at https://www.bilibili.com/video/BV19U4y1q781?spm_id_from=333.999.0.0.

Keywords Multisensor fusion · GNSS · Visual · Outlier exclusion

1 Introduction

Localization is a key part of autonomous vehicle applications, such as autonomous driving, urban air mobility, virtual reality and augmented reality.

Traditional positioning algorithms used for vehicles use an integrated system combining an inertial navigation system (INS) and global navigation satellites system (GNSS).

On the other hand, visual odometry, lidar odometry or radar odometry have been popular research topics in recent decades. Normally, such a technique was proposed for applications in simple scenarios. And it may not be robust enough for practical application, where there may be challenges such as urban canyon scene, indoor-outdoor transition and dynamic environment.

B. Zhang (✉) · X. Wan · L. Shu
Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, No. 9, Dengzhuang South Road, Haidian District, Beijing, China
e-mail: zhangbh@csu.ac.cn

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_14

To improve the robustness of the state estimation system, there are mainly two ways.

The first one is multisensor fusion. These sensor fusion methods have been popular in recent years for the complementary properties provided by heterogeneous sensors: the camera provides rich visual information with a low cost in good lighting condition, the inertial measurement unit (IMU) offers high frequency and outlier-free attitude and acceleration measurement, but with accumulated error the GNSS provides a drift-free localization in the global frame in opening areas with enough satellites in sight, etc.

The second way is data handling methods. For the camera, there may be low quality data caused by disturbance from dynamic object, textureless environment or a poor lighting condition. For the GNSS, there has always been a major data problem caused by the multipath effect, particularly in a deep urban area. The data handling methods can be divided into two classes: data weighting and data exclusion methods.

The data weighting methods calculate weights for different sensors. Sensor's measurements with a large deviation from the model used would have a reduced weight on the final result. Some popular data weighting methods are: the maximum likelihood estimators (or M-Estimators) [1], Switchable Constraints [2], dynamic covariance scaling [3] and max mixtures [4].

The data exclusion methods try to find a reliable subcollection of all the measurements, in other words, eliminating the outliers. Some popular data exclusion techniques are: rANdom sAmple consensus (RANSAC) [5], realizing, reversing, recovering (RRR) [6] l1 relaxation [7] and receiver autonomous integrity monitoring (RAIM) [8–10].

Generally, the data weighting methods mainly focus on improving the estimation accuracy by finding the more reliable sensor's data; while the data exclusion methods focus on preventing the crash of the system by kicking out the bad sensor's data, which make it a more fundamental task.

At last, for the data handling methods in the multisensor fusion frameworks, there have been various kind of techniques in previous works: Chiang et al. [11] use the comparatively short-period INS navigation solution to monitor the lidar odometry result and then use the lidar odometry result to monitor the comparatively long-period GNSS position and velocity results. Liang et al. [12] use separate outliers exclusion methods in radar, lidar and camera process in a loosely coupled error-state extended Kalman Filter-based framework. Chu et al. [13] implemented a multi-layer RANSAC scheme in the visual data processing in a tightly coupled fusion EKF-based framework. Meng et al. [14] proposed a multiconstraint fault-detection methods to suppress the GNSS outliers and false curves or points of the lidar by using the RANSAC algorithm. Santamaria-Navarro et al. [15] try to achieve robustness through redundant, parallel sensors and state estimators, then generate a smooth state estimation by multiplexing the separated estimators, confidence tests for data quality and algorithm health were also performed.

In this paper, we firstly study the data exclusion methods on the visual and GNSS data in a tightly couple, optimization-based state estimation framework with three type of sensors (IMU, monocular camera and GNSS receiver). Secondly, we try to

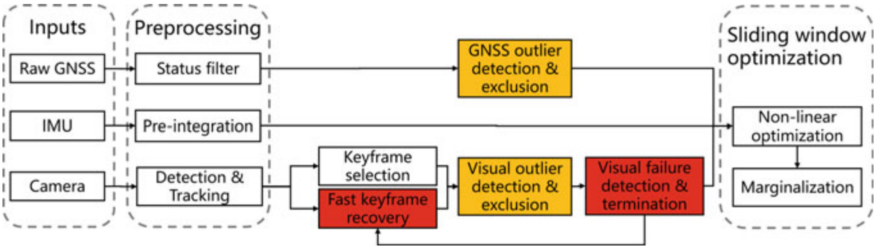


Fig. 1 Diagram of our major contributions in the odometry system

find some countermeasures in the situation that one kind of sensor's data, the visual data in our case, is completely failed.

Figure 1 shows the three major parts of our work in the odometry system: all these modules are added between the front end and back end of the odometry system.

Our contributions are

- A sliding window-based RAIM is proposed for the GNSS data and re-projection error-based outlier exclusion for the visual data in a tightly couple optimization-based GNSS/visual/inertial odometry;
- We keep the state estimation stable when complete failure occurs in the visual data by visual termination and keyframe fast recovery;
- Evaluation test of the proposed methods in real-world environments.

2 Methodology

2.1 GNSS Measurement Outlier Detection and Exclusion

For GNSS data, the RAIM method had been broadly utilized in the practical application. We extended this method in a sliding window optimization-based odometry. This idea is introduced below:

The basic GNSS measurements relationship is described by a linear equation in the form of

$$\Delta\rho = G\Delta x + \Delta\epsilon \tag{1}$$

where n is the number of redundant measurements; Δx is the 3×1 vector of the receiver's true position plus the clock bias of the receiver (as a 4×1 vector); $\Delta\rho$ stands for the difference between the actual measured range (namely, pseudorange) and the predicted range from the receiver's nominal position with the clock bias (as an $n \times 1$ vector); $\Delta\epsilon$ is the measurement error from the receiver noise, various interferences in signal propagation, satellite position errors and satellite clock error(as an $n \times 1$ vector) and G is the linear coefficient matrix between Δx and $\Delta\rho$ (as an $n \times 4$ matrix).

In most references, the estimation uses the least squares solution to the measurement equations at a single time, which is called single-point solution. To improve the navigation precision, we follow GVINS [16] to use an iterated optimization method to get the solution, which may take relatively more computation time.

While the above navigation solution is obtained in the Earth-Centered Earth-Fixed (ECEF) frame, we will calculate the GNSS measurements residual in the Earth-Centered Inertial (ECI) frame. The GNSS measurements are time-stamped by the receiver. Defining the ECI frame to be coincident with the ECEF frame when the signal arrives at the receiver, we have

$$p_{\mathrm{r}}^{\mathrm{E}} = p_{\mathrm{r}}^{\mathrm{e}} \tag{2}$$

where $p_{\mathrm{r}}^{\mathrm{E}}$ is the receiver position in the ECI frame; $p_{\mathrm{r}}^{\mathrm{e}}$ is the receiver position in the ECEF frame, corresponding to the first three elements of Δx in (1).

The satellite's position $p_{\mathrm{s}}^{\mathrm{e}'}$ in ECEF frame when the signal is transmitted can be calculated by the satellite's ephemeris and the pseudorange measurement. The satellite's position in the ECI frame as a result of Earth's rotation became

$$p_{\mathrm{s}}^{\mathrm{E}} = R_z(-\omega_{\mathrm{E}} t_{\mathrm{f}})\, p_{\mathrm{s}}^{\mathrm{e}'} \tag{3}$$

where R_z stands for a rotation in the z axis of the ECI frame, ω_{E} is the angular velocity of the Earth rotation and t_{f} is the transmission time of the GNSS signal.

The original residual of a single pseudorange measured in t_k with respect to the n GNSS measurements can be formulated as

$$r_k^{\mathrm{p}} = \left\| p_{\mathrm{s}}^{\mathrm{E}} - p_{rk}^{\mathrm{E}} \right\| + c\left(\delta t_k - \Delta t^{sj}\right) + T_{rk}^{sj} + I_{rk}^{sj} - \tilde{P}_{rk}^{sj} \tag{4}$$

where r_k is the GNSS receiver at time t_k, c is the speed of light, δt_k is the clock bias corresponding to the four elements of Δx in (1), t^{sj} is the clock error of satellite j, T_{rk}^{sj} and I_{rk}^{sj} are the tropospheric and ionospheric delay of the signal from satellite j, and $\tilde{P}_{rk}^{sj}$ stands for the pseudorange measured by the GNSS receiver.

To take the satellite ephemeris, satellite elevation angle and the pseudorange measurements error into consideration, we simply define

$$r_k = \frac{r_k^{\mathrm{p}} \sin^2 \mathrm{el}}{(\mathrm{ura} - 1)\sigma_{\mathrm{psr}}} \tag{5}$$

where el stands for the elevation angle of the GNSS satellite, *ura* is the satellite signal accuracy index from the ephemeris and σ_{psr} is the standard deviation of the pseudorange measurements from the GNSS receiver.

For more detail, please refer to GVINS [16] and Spilker et al. [9].

Normally, the above residual is calculated within the current epoch of the GNSS measurements, and pseudoranges with residual higher than a predefined threshold, (namely GNSS outlier threshold) are considered to be outliers of the GNSS measurements. However, the satellite number could become very poor in the urban canyon environment, which will make the residual in (5) less effective to help us finding the GNSS signal outlier.

When using an optimization-based odometry, we find the past GNSS measurements in the sliding window could be very helpful in calculating the residual of the current measurements (which is the last measurement in the sliding window). Although these old GNSS measurements were caught in previous locations of the vehicle. The location difference is relatively small compared with the bad pseudorange measurements as the window size is limited. However, the increased measurements number will make the residual much more credible.

2.2 Visual Measurement Outlier Detection and Exclusion

For visual data, RANSAC has always been the most popular method in outlier detection and exclusion. However, in this work, we simply use the reprojection error defined in [17] to find the visual measurement outlier.

Let the sliding window size of visual frame to be m. Consider a visual feature l that is both observed in the $m-1$th frame and the mth frame.

The feature location in the unit plane ($z = 1$) of the camera frame is

$$\overline{P}_l^m = \pi_c^{-1}\left(\begin{bmatrix} u_l^m \\ v_l^m \end{bmatrix}\right) \tag{6}$$

where $\left[u_l^m, v_l^m\right]^{\mathrm{T}}$ is pixel location of the lth feature that found by the optical flow tracking in the mth frame, π_c^{-1} is the back-projection function which turns a pixel location into a unit plane vector in the camera frame.

We use the following equation to transfer the pixel location of the lth feature in the m-1th frame to the mth frame:

$$\overline{P}_l^m = T_b^c T_w^{b_m} T_{b_{m-1}}^w T_c^b \frac{1}{\lambda l} \pi_c^{-1}\left(\begin{bmatrix} u_l^{m-1} \\ v_l^{m-1} \end{bmatrix}\right) \tag{7}$$

where λl is the inverse depth of the lth feature, T_c^b is the transition matrix from the camera coordinate system to the IMU coordinate system, T_b^c is the inverse of T_c^b, $T_{b_{m-1}}^w$ is the transition matrix from the $m-1$th camera frame to the world coordinate system and $T_{b_{m-1}}^w$ is the transition matrix from the world coordinate system to the mth camera frame. We get T_b^c by calibrating the extrinsics of the camera and the IMU. And λl, $T_{b_{m-1}}^w$, $T_w^{b_m}$ are part of the state variables to be solved in the odometry.

In the end, the visual residual for the feature l in the last frame of the sliding window is defined as:

$$r_l^{\mathrm{v}} = \left\| \overline{P}_l^{m-1} - \frac{P_l^{m-1}}{P_l^{m-1}} \right\| \tag{8}$$

Features with visual residual higher than a predefined threshold (namely, visual outlier threshold) are considered to be outliers of the visual measurements.

2.3 Visual Data Failure Handling

In some cases, external interference could make the visual data a complete failure. For instance, when an autonomous driving car stop at a traffic light, people who is walking down the pedestrian crosswalk may completely block the view of the camera on the car. This may eliminate most of the feature points in the visual front end, and the remaining feature points may have terrible optical flow tracking results in the coming new frame.

To cope with this situation, we propose a simple principle called visual termination: after the visual measurement outlier detection and exclusion module in the system, whenever the number of the remaining feature points is lower than a predefined threshold (namely, visual termination threshold), the visual data is cut off from the back end optimization process.

Because only the IMU and the GNSS data are fed into the optimization in this visual termination state, noise in the result of the odometry will be increased by the GNSS data, especially in the vertical direction. So, the visual data should be recovered as soon as the external interference has gone.

To recover the visual data, a keyframe should be selected at the first place in a feature-based visual odometry. Popular VO system such as VINS-mono [17] or ORB-SLAM2 [18] selects keyframe by rules like:

- Current frame shares few feature points with the latest keyframe, or the average parallax between current and the latest keyframe is over the limit;
- There has been enough time or frames passed between current frame and the latest keyframe.

In order to recover the visual data as soon as possible in the visual termination, we propose another simple principle called keyframe fast recovery: in the middle of the visual termination state, whenever the feature points number in a frame is higher than a predefined threshold (namely, keyframe recovery threshold), this current frame will be selected as a keyframe and sent into the optimization.

3 Field Testing

To verify the methodologies we mentioned above, we run our odometry in many different scenes and find three typical ones as our final tests:

Test 1 for GNSS outlier exclusion: A 160 m straight line trajectory in an office park surrounded by up to 40-m office buildings, as shown in Fig. 2;
Test 2 for visual outlier exclusion: A 245 m circular line trajectory surrounded by 15-m residential buildings, as shown in Fig. 3;

Fig. 2 Trajectory and scene for test 1

Fig. 3 Trajectory and scene for test 2

Test 3 for visual data failure handling: About 10 m forward to fully initialize the system, then hold still by the roadside, waiting for pedestrian or vehicle to pass in front of the camera.

All the tests were carried out by walking while holding an Intel Realsense D435i camera [19] with camera and IMU inside, an U-blox ZED-F9P GNSS receiver [20] and a laptop.

For all tests, we set the GNSS outlier threshold = 100, the visual outlier threshold = 20, visual termination threshold = 10 and the keyframe recovery threshold = 20.

The local East-North-Up coordinate system is defined as: the origin is at the point when the GNSS data is firstly fused into the odometry, the X, Y, Z axis are pointing to the east, north, up direction, respectively.

In the test results, we use the following abbreviations for the odometry results:

GVINS_default: The original GVINS [16] odometry;
GVINS_GNSS-OE: Modified version of GVINS with the GNSS outlier exclusion method;
GVINS_Visual-OE: Modified version of GVINS with the visual outlier exclusion method;
GVINS_OE: Modified version of GVINS with both the GNSS and visual outlier exclusion method;
GVINS_FastRecovery: Modified version of GVINS with the visual data failure handling method.

Finally, we analyze the increase of computation time due to the outlier detection and exclusion modules in the odometry.

3.1 Test 1: GNSS Outlier Exclusion

Test 1 is performed in a typical urban canyon environment with serious multipath effect and limited satellite number.

The absolute trajectory error (ATE) of different estimator is shown in Fig. 4.

For GVINS_default: huge translation error accumulated in the Z direction (downward).

For GVINS_GNSS-OE: while there is still error in all coordinate axis, error in the vertical direction was greatly removed.

3.2 Test 2: Visual Outlier Exclusion

In test 2, the GNSS satellite signal is still not very good because our path is too close to the buildings, and the visual feature points residuals appear to be large due to the fast motion of the sensors. These residuals are shown in Fig. 5.

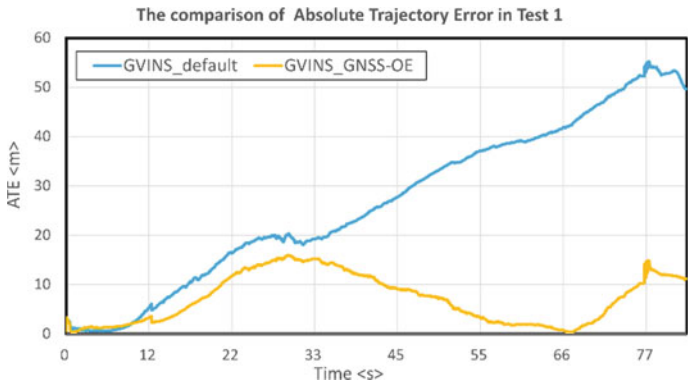


Fig. 4 Absolute trajectory error of GVINS_default and GVINS_GNSS-OE in test 1

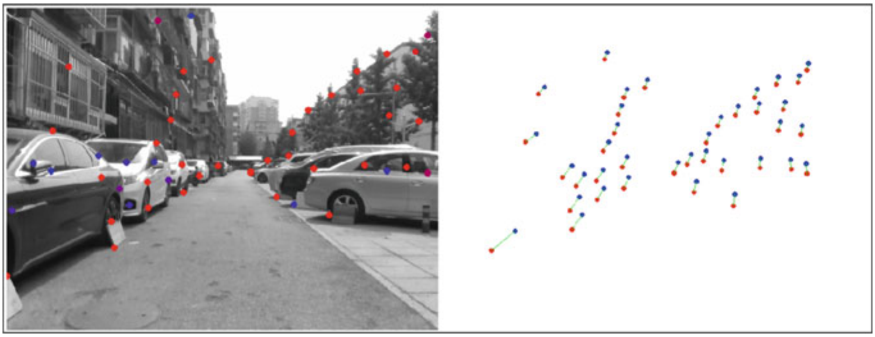


Fig. 5 Visual feature points (left) and residual plot in two contiguous frames (right), the red points stand for the feature points in the m-1th frame in the sliding window, the blue points stand for the feature points in the mth frame and the green lines stand for the visual residuals of the two frames

The absolute trajectory error (ATE) of different estimator is shown in Fig. 6.

For GVINS_default: wrong initial direction of the odometry leads to big translation error in the X-Y plane at the beginning, then the drift was corrected by the GNSS data. But the error grew again at the sharp turning.

For GVINS_Visual-OE: initial direction was greatly corrected and translation error was relatively small in the whole process.

3.3 Test 3: Visual Data Failure Handling

A cycling person sweeps through the camera view at 11 s of the test, then followed by a car at 18 s. They kill all the feature points in the odometry front end, as shown in Fig. 7.

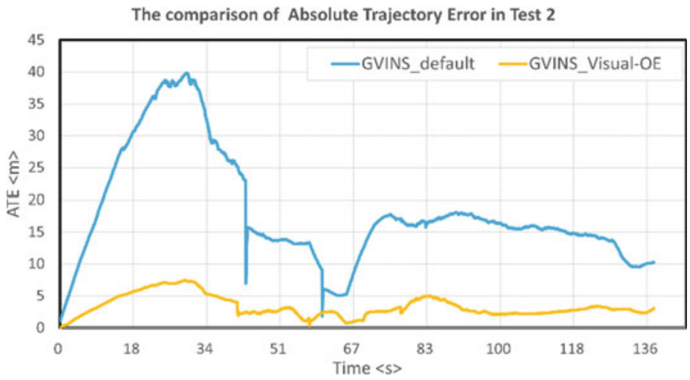


Fig. 6 Absolute trajectory error of GVINS_default and GVINS_Visual-OE in test 2

Fig. 7 Visual feature points fail because of a moving object in the camera view, there is nothing in the residual plot compared with Fig. 5

The absolute trajectory error (ATE) of different estimators is shown in Fig. 8.

For GVINS_default: after the sweeping, the odometry diverge slowly at first, then in tens of second there is a large error in the Z direction(downward); finally new feature points were generated and the odometry recovers to normal levels.

For GVINS_OE: generally the same as the default version except for some small trajectory differences.

For GVINS_FastRecovery: the divergence never occurs and the trajectory holds for the whole process.

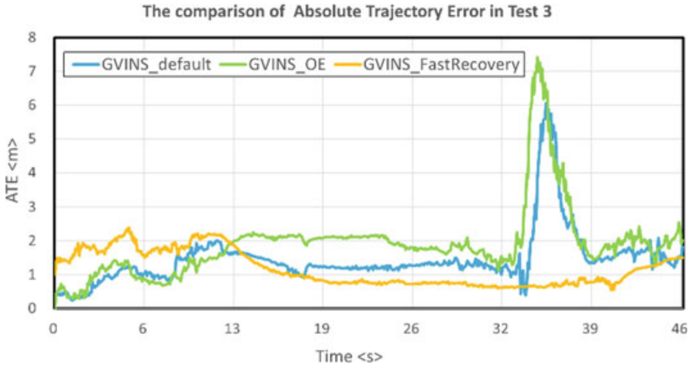


Fig. 8 Absolute trajectory error of GVINS_default, GVINS_OE and GVINS_FastRecovery in test 3

3.4 Timing Statistics

We run the data of the tests on a laptop with Intel i5-8300H CPU running at 2.30 GHz. Timing statistics of the outlier detection and exclusion modules in the odometry backend thread are given in Table 1.

While other modules of the odometry (pre-integration of the IMU data, optimization, marginalization, etc.) are not the focus points of this paper, we can see from Table 1 that the position determination process by pseudorange measurement takes up the overwhelming majority of the computation resource in the outlier detection and exclusion module. This process is basically the iterations of Eq. (1) to calculate the receiver position. We also find that the calculation time of the receiver position increases monotonically with the number of available GNSS satellites, as shown in Fig. 9.

As our GNSS outlier detection and exclusion module is built to handle the multipath effect in urban canyon scenes, where the available GNSS satellites number is much lower than that in open spaces, real-time performance can be guaranteed, so the increase in computation time is acceptable.

Table 1 Timing statistics of the odometry (with seven GNSS satellites available)

Modules	Time (ms)
GNSS_OE: Position determination by pseudorange measurement	15.463
GNSS_OE: Other parts of the GNSS outlier exclusion module	0.059
Visual_OE	0.014
Other modules of the odometry	54.054
Total	69.589

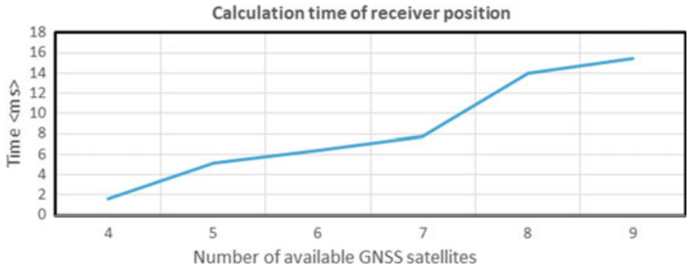


Fig. 9 Calculation time of the receiver position

4 Conclusion

In this paper, we propose outlier detection and exclusion methods for the visual data and GNSS data in a tightly couple, optimization-based GNSS/visual/inertial odometry. We also propose a visual termination and keyframe fast recovery strategy to handle the complete failure of the visual data. Real-world tests have shown the effectiveness of our methods in improving the robustness of the odometry and increase of computation time due to the methods are analyzed. Although the job is built based on an optimization odometry, we believe these methods also work in a Kalman Filter-based odometry.

In future work, other types of sensors, such as lidar and wheel speedometer, will be included. We will try to find the suitable outlier detection and exclusion methodology for the new sensors' data and we will also try to find whether it would be necessary to design a termination strategy for these new sensors in their related corner cases.

References

1. Zhang Z (1997) Parameter estimation techniques: a tutorial with application to conic fitting. Image Vis Comput 15(1):59–76
2. Sunderhauf N (2012) Robust optimization for simultaneous localization and mapping. Ph.D. dissertation, Technischen Universitat Chemnitz
3. Agarwal P, Tipaldi GD, Spinello L, Stachniss C, Burgard W (2013) Robust map optimization using dynamic covariance scaling. In: 2013 IEEE international conference on robotics and automation, 06–10 May 2013
4. Olson E, Agarwal P (2013) Inference on networks of mixtures for robustrobot mapping. Int J Robot Res 32(7):826–840
5. Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM 24(6):381–395
6. Latif Y, Cadena C, Neira J (2012) Realizing, reversing, recovering: Incremental robust loop closing over time using the iRRR algorithm. In: 2012 IEEE/RSJ international conference on intelligent robots and systems. IEEE, pp 4211–4217
7. Carlone L, Censi A, Dellaert F (2014) Selecting good measurements via l1 relaxation: a convex approach for robust estimation over graphs. In: 2014 IEEE/RSJ international conference on intelligent robots and systems. IEEE, pp 2667–2674

8. Grover Brown R (1992) A baseline GPS RAIM Scheme and a Note on the Equivalence of three RAIM methods. J Inst Navig 39(3):301–316
9. Spilker JJ Jr, Axelrad P, Parkinson BW, Enge P (1996) Global positioning system: theory and applications, 2-volume sets. American Institute of Aeronautics and Astronautics, Inc., Reston, pp 143–165
10. Hewitson S, Wang J (2006) GNSS receiver autonomous integrity monitoring (RAIM) performance analysis. GPS Solut 10:155–170
11. Chiang K-W, Tsai -J, Li Y-H, Li Y, El-Sheimy N (2020) Navigation engine design for automated driving using INS/GNSS/3D LiDAR-SLAM and integrity assessment. Remote Sens 12
12. Liang Y, Muller S, Schwendner D, Roll D, Ganesc D, Schaffer I (2020) A scalable framework for robust vehicle state estimation with a fusion of a low-cost IMU, the GNSS, Radar, a Camera and Lidar. In: 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS), Las Vegas, NV, USA, 25–29 Oct 2020
13. Chu T, Guo N, Backén S, Akos D (2012) Monocular camera/IMU/GNSS integration for ground vehicle navigation in challenging GNSS environments. Sensors 3162–3185
14. Meng X, Wang H, Liu B (2017) A robust vehicle localization approach based on GNSS/IMU/DMI/LiDAR sensor fusion for autonomous vehicles. Sensors 17
15. Santamaria-Navarro A, Thakker R, Fan DD, Morrell B, Agha-mohammadi A (2020) Towards resilient autonomous navigation of drones. arXiv:2008.09679v1 [cs.RO] 21 Aug 2020.
16. Cao S, Lu X, Shen S (2022) GVINS: tightly coupled GNSS–visual–inertial fusion for smooth and consistent state estimation. IEEE Trans Robot
17. Qin T, Li P, Shen S (2018) VINS-mono: a robust and versatile monocular visual-inertial state estimator. IEEE Trans Robot 34(4)
18. Mur-Artal R, Tardós JD (2017) ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Trans Robot 33(5)
19. Intel RealSens Depth Camera D435i. https://www.intelrealsense.com/depth-camera-d435i/
20. ZED-F9P module. https://www.u-blox.com/en/product/zed-f9p-module

Clinical Nurses Before and After Simulated Postoperative Delirium Using a VR Device Characteristics of Postoperative Delirium Imagery

Jumpei Matsuura, Yoshitatsu Mori, Takahiro Kunii, and Hiroshi Noborio

Abstract The purpose of this chapter was to have clinical nurses simulate the hallucinations and auditory hallucinations experienced by patients with delirium using a VR device, and to clarify the changes in the clinical nurses' perception of delirium before and after the experience. Meta Quest was used for the Head-Mounted Display (HMD), which is a VR device, and the VR content for the simulated experience of delirium used in the device was created independently using Unity. The created VR content reproduced the hallucinations experienced by patients with postoperative delirium in the ICU at night. The duration was 12 minutes. The contents were set to change at 2-minute intervals, with scenes of cockroaches appearing on the ceiling, the ceiling closing in, a person in protective clothing appearing, and soldiers attacking.

Keywords VR device · Postoperative delirium · Nursing

1 Problem Statement

Postoperative delirium is transient disturbance of consciousness that commonly occurs in elderly patients undergoing surgery under general anesthesia. There are three factors that contribute to the development of postoperative delirium: preparatory factors, direct factors, and precipitating factors. Preparative factors include aging, dementia, and cerebrovascular disease. Direct factors include drug addiction,

J. Matsuura (✉)
Nara Gakuen University, 3-15-1 Naka Tomigaoka, Nara City, Nara Prefecture, Japan
e-mail: jmatsuura@nara-su.ac.jp

Y. Mori · H. Noborio
Osaka Electro Communication University, Osaka, Japan

T. Kunii
Kashina System Co, Hikone, Japan

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_15

metabolic disease, and alcohol withdrawal. Inducing factors include psychological and social stress, sleep disturbances, sensory deprivation, and physical restraints [1].

Postoperative delirium is a type of delirium and refers to the appearance of delirium in patients after surgery [2]. The main symptoms of delirium include hallucinations, such as seeing small insects on the ceiling that are not supposed to be there, and hallucinations in which persons who are not there actually appear to be there.

Postoperative delirium is known to cause many difficulties for nurses in providing nursing care. It has been reported that 30% of patients with delirium remember their experiences [3]. Patients with delirium may be traumatized by the fearful experiences they have had, or they may feel remorse for their own verbal abuse or violence.

In a previous study, researchers conducted a study on nursing students using a VR device similar to the one used in this study. The results showed that there was no significant difference in both the amount and duration of speech before and after the experience of VR content (hereinafter referred to as "VR experience"). However, after the VR experience, more statements were heard in which the students felt the need to be close to the patient. The results also showed a better understanding of the inner life of patients who were experiencing fear due to delirium.

The purpose of this study was to reproduce the hallucinations experienced by patients with delirium using a VR device, and to have nurses with more than 10 years of clinical experience in the surgical field simulate these hallucinations, and to clarify the changes in perception of delirium that occur before and after these experiences.

2 Approach

Eleven nurses with more than 10 years of clinical experience in the surgical field were subjects of this study, and the HMD for the VR experience was the Meta Quest. The subjects were asked to wear the HMD and lie on the bed for the VR experience to reproduce the same situation as a patient with postoperative delirium (see Fig. 1).

Fig. 1 Experimental scene

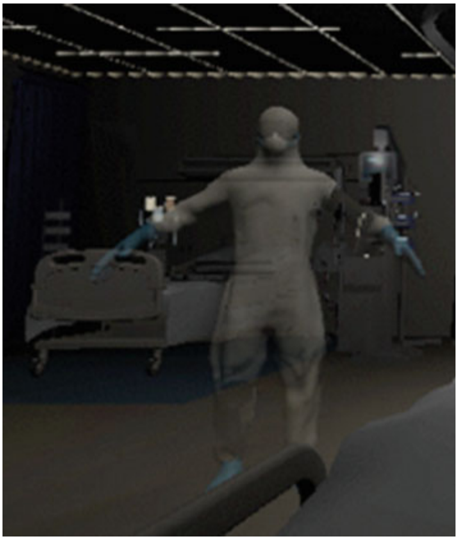


Fig. 2 A figure in protective clothing appears

VR contents to simulate the delirium experience (see Fig. 2) were created using Unity Approach.

The target nurses were asked to answer two questions before and after the experience. Question 1 was about the image of a patient with delirium, and Question 2 was about the setting of the onset of delirium. A verbatim transcript was made from the interviews.

The analysis was conducted on speech duration and speech volume, and compared before and after VR viewing. Analysis methods were Wilcoxon rank sum test and Hierarchical cluster analysis. SPSS Ver. 26 was used as the statistical processing software. KH coder was used as the metric text analysis software.

3 Results

The speaking time during the interview with the nurses was as follows. 1.5 minutes before and 2.39 minutes after the experience for Question 1, and 1.62 minutes before and 1.7 minutes after the experience for Question 2. There were no significant differences between the pre- and post-experience times for questions 1 (0.084) and 2 (0.753) (Tables 1 and 2).

Table 1 Speech time

Question 1		
Before (SD)	After (SD)	*p*
1.5 (1.063)	2.39 (2.238)	0.084

Table 2 Speech time

Question 2		
Before (SD)	After (SD)	*p*
1.62 (1.151)	1.7 (1.407)	0.753

The speech volume was as follows. 294.5 words before the experience for question 1 and 448.2 words after the experience. The number of words spoken before and after the experience with Question 2 was 344.6 and 332.8, respectively. There was no significant difference between the pre- and post-experience of question 1 (0.345) and question 2 (0.917) (Tables 3 and 4).

Text mining analysis revealed that the number of words spoken by the nurses was 294.5 before the experience of Question 1. The main words spoken included "Image" (5.4%), "Delirium" (4.7%), and "Patient" (4.4%) (Table 5).

After the experience of Question 1, 448.2 words were used. The main words were "See" (6.2%), "Feel" (6.0%), and "Think" (6.0%) (Table 6).

Before the experience of Question 2, 344.6 words were used. The main words were "Patient" (9.2%), "Angry" (4.6%), and "Wonder" (4.0%) (Table 7).

After the experience of question 2, 332.8 words were used. The main words were "Think" (9.9%), "Feel" (7.5%), and "Patient" (6.0%) (Table 8).

Table 3 Speech volume

Question 1		
Before (SD)	After (SD)	*p*
294. 5 (284.19)	448.2 (507.89)	0.345

Table 4 Speech volume

Question 2		
Before (SD)	After (SD)	*p*
344.6 (204.87)	332.8 (204.19)	0.917

Table 5 Q1 before drawing volume

Drawing	Frequency	Drawing	Frequency
Image	16	Understand	5
Delirium	14	Nurse	5
Patient	13	Medicine	4
Imagine	8	Person	4
People	7	Possible	4
Wonder	7	Medicine	4

Table 6 Q1 after drawing volume

Drawing	Frequency	Drawing	Frequency
See	28	Guess	7
Feel	27	Fear	6
Think	27	Image	6
Patient	24	Bother	5
Sound	16	Dark	5
Wonder	15	Eye	5
Situation	11	People	5
Monitor	11	Scare	5
Delirium	9	Sleep	5
Anxious	7	Time	5

Table 7 Q2 before drawing volume

Drawing	Frequency	Drawing	Frequency
Patient	32	Restrain	5
Angry	21	Delirium	4
Wonder	14	Honest	4
Help	10	Sad	4
Feel	9	Shift	4
Work	9	Emotion	3
Something	8	Frustrated	3
Time	6	Patients	3
Calm	5	Priority	3
Guess	5	Restrain	3
Person	5	Sorry	3

Table 8 Q2 after drawing volume

Drawing	Frequency	Drawing	Frequency
Think	33	Time	6
Feel	25	Bad	5
Patient	20	Cancel	5
A	12	Happen	5
Help	10	Understand	5
Wonder	10	World	5
Feeling	8	Experience	4
Something	8	Information	4
Nurse	7	Medication	4
See	7	Person	4
Frustrated	6	Shift	4

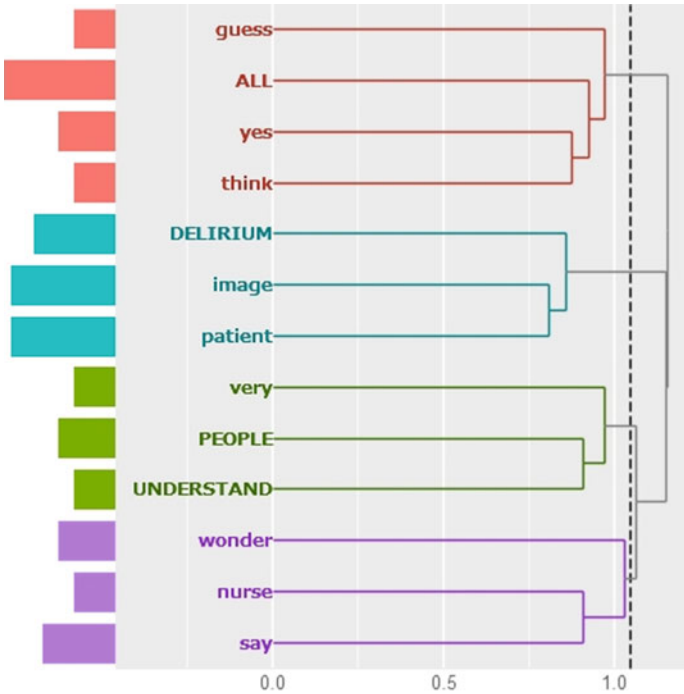


Fig. 3 Question 1 before experiencing VR content cluster analysis results

The results of hierarchical cluster analysis showed that respondents were classified into four to six categories before and after each experience (see Figs. 3, 4, 5, and 6) for questions 1 and 2.

4 Discussion

In recent years, a number of contents that allow users to experience VR have been tackled in the field of medicine. Specifically, they include surgical simulations [4–6].

The results of interviews with nurses who experienced VR were analyzed using quantitative evaluation and text mining.

In the quantitative evaluation, there were no significant differences in both speech duration and speech volume. However, with regard to speaking time and volume, question 1 asked respondents to speak freely about the image they had of the delirium patient. In question 1, there was an increase in both speaking time and volume before and after the experience. We speculate that this may be an indication that the sympathetic nervous system is slightly predominant during the VR experience. This result is consistent with that of Yamashita [7].

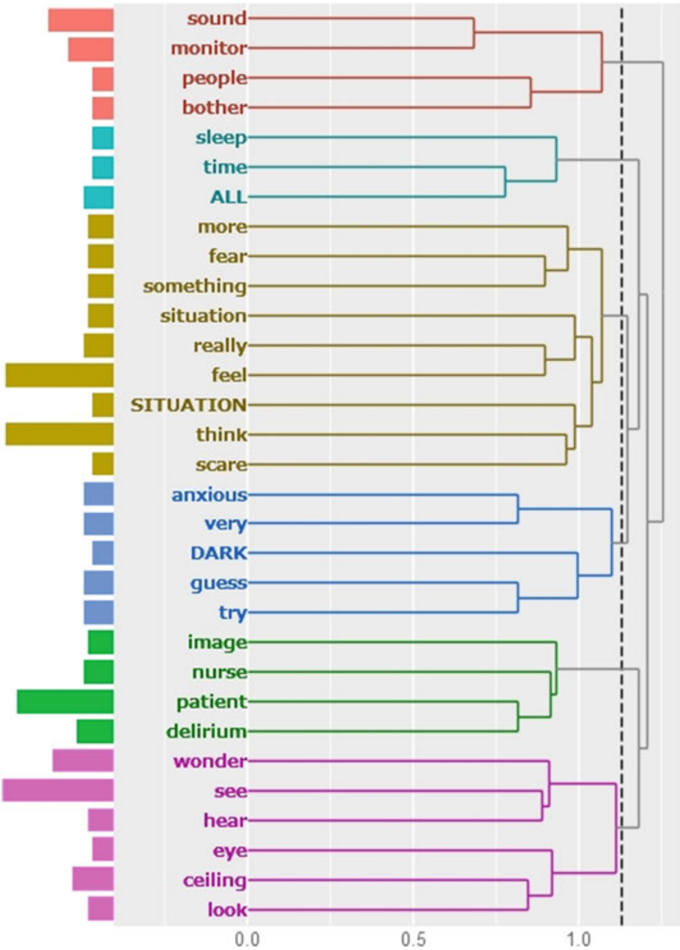


Fig. 4 Question 1 after experiencing VR content cluster analysis results

Question 2 is a situational question. Specifically, "You have an important appointment after your night shift ends. However, because of the patient's delirium, you cannot finish your work on time and you have to work overtime. If the patient is the cause, how do you feel about the patient?" The question was, "If the patient is the cause of the problem, how do you feel about the patient?"

Regarding question 2, before and after the experience, conversely, the volume of speech decreased after the experience. This is because many of the patients before the experience prioritized the convenience of the nurses, whereas after the experience, many of them said that they wanted to be there for the patients rather than for the nurses' convenience. We speculate that the nurses' understanding of the painful inner life that delirium patients experience may have led to the decrease in the amount of speech.

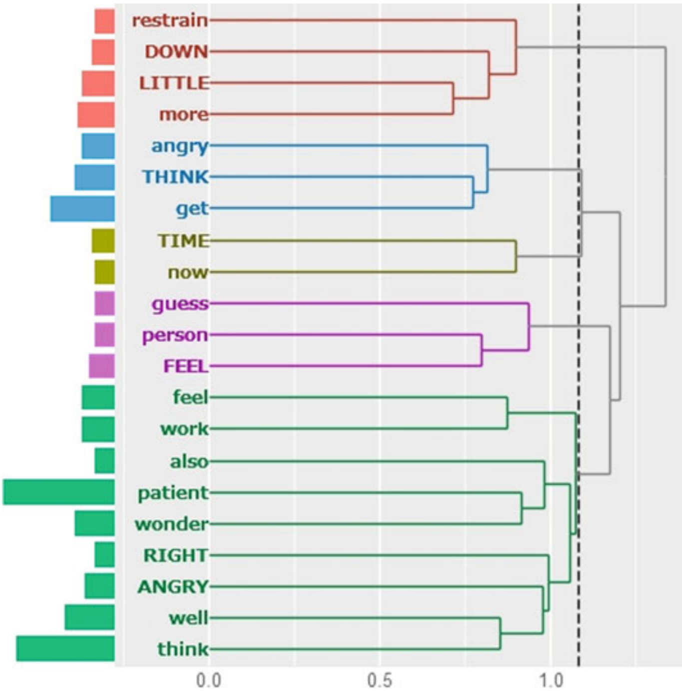


Fig. 5 Question 2 before experiencing VR content cluster analysis results

The results of the text mining revealed that before experience of question 1, understanding the patient, etc. were extracted. After the experience, words such as the sound of the monitor and fear were extracted. These words did not appear before the experience, and we believe that they appeared only because the participants experienced the delirium simulation through the VR content. From this, we infer that the VR experience may have led to a better understanding of inner life of patients with delirium.

Before and after the experience of Question 2, the characteristics of the patients were extracted, such as stress build-up due to anger toward the patient who developed delirium. However, we speculate that the nurses' simulated experience of the strange experiences experienced by the delirium patient may have led to a better understanding of the patient's inner world.

The hallucinations of delirium patients cannot be actually seen by anyone other than the patients themselves. However, by using VR contents, it is possible to visualize them. Visualization makes it easier to imagine and deepen understanding [8]. We believe that simulating the experience of a patient with delirium will lead to better understanding of the patient and provide nursing care that is close to the patient.

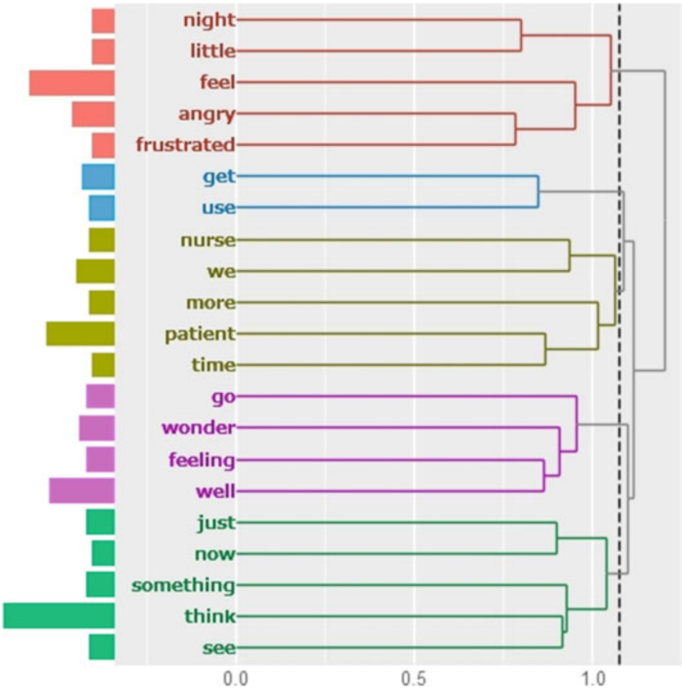


Fig. 6 Cluster analysis results after experiencing Question 2 VR content

5 Conclusion

The results of clinical nurses' simulated experience of postoperative delirium using VR content suggest that understanding inner life of patients with postoperative delirium, such as their fear, can help nurses provide nursing care that is more attentive to patients. The results suggest the need for further clinical education for nurses using VR content.

6 Future Work

Eleven clinical nurses were included in this study. This number is not considered large by any means. Therefore, we feel that it is necessary to increase the number of subjects for future studies.

References

1. Lipowski ZJ (1990) Delirium. Acute confusional states. Oxford University Press, New York, pp 54–70
2. Takeuchi M, Yamamoto A, Shimada Y et al (2009) Frequency and characteristics of delirium in hospitalized patients toward devising a delirium risk factor check sheet. Hamamatsu Rosai Hospital Academic Annual Report, pp 30–32
3. Inouye SK (2001) Nurses' recognition of delirium and its symptoms comparison of nurse and researcher ratings. Arch Intern Med 161:2467–2473
4. Sugimoto M, Yasuda H, Koda K, Suzuki M, Yamazaki M et al (2010) Image overlay navigation by markerless surface registration in gastrointestinal, hepatobiliary and pancreatic surgery. J Hepatobil Pancreat Sci 17(5):629–636
5. Hayashi Y, Misawa K, Hawkes DJ, Mori K (2016) Progressive internal landmark registration for surgical navigation in laparoscopic gastrectomy for gastric cancer. Int J Comput Assist Radiol Surg 11(5):837–845
6. Deng W, Li F, Wang M, Song Z (2014) Easy-to-use augmented reality neuronavigation using a wireless tablet PC. Stereotact Funct Neurosurg 92(1):17–24
7. Yamashita Y (2020) The quintessence, 0286-407X. 39(6):1412–1417
8. Noguchi Y, Ito T, Yokota M (2019) Application of Virtual Reality for understanding the living environment. Occupational Therapy 0289–4920(6):736–740

Modeling and Simulation of a Frequency Reconfigurable Circular Microstrip Antenna Using PIN Diodes

Ashrf Aoad

Abstract In this study, a frequency reconfigurable circular microstrip antenna for wireless communications has been presented. PIN diodes are integrated to obtain multiple-band operation. The operating frequency can be tuned to other frequencies by switching ON/OFF states of the diodes. Many approaches have been explored, which provide multiple wideband operations when integrated with identical microstrips to be regularly controlled. The first one simulates the main circular part without any switching state, then steeply different switching states start. The operating frequencies are between 5 and 15 GHz. It operates exactly in the most tested and demanded dual five-generation (5G) and single industrial, scientific, and medical (ISM) 5.2 GHz bands. The obtained bands are considered for different wireless communication application uses. All results are performed with the method of finite integration technique (FIT).

Keywords Multiple wideband · Reconfigurable circular antenna · PIN diodes · Switching · Wireless communication

1 Introduction

Antennas are necessary and critical components of wireless communication systems. Arguably, different types of antennas have been developed and shown in the communication engineering market during the past years, (e.g., dipoles, microstrip, loop antennas, and frequency-independent antennas). Reconfigurable antennas have the capabilities and potential to develop new possibilities for communication systems' performance. Frequency reconfigurable antennas have become important and started many years ago in cellular radio communication, radar systems, airplane, satellite, mobile and microwave link networks [1–3]. In some communication systems such as mobile and satellite, reconfigurable antennas are useful for supporting a large number

A. Aoad (✉)
Istanbul Sabahattin Zaim University, Istanbul, Turkey
e-mail: esref.osmanlioglu@izu.edu.tr

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_16

of universal applications [(e.g., Wi-Fi, UMTS-3G, Bluetooth, WiMAX, and DSRC standards)] to decrease strong signal coupling and interference in the microwave environment. Thereby, it is a notable need for adapting new technology of reconfigurable antennas. They present the most compact solutions and additional capabilities for communication application requirements by controlling the same reconfigurable antenna for fundamental operating characteristics such as frequencies, s-parameters, the direction of radiation pattern, polarization, antenna directivity (D), gain (G), and efficiency (η) that relates in the antenna according to ($G = \eta D$). In general, reconfiguring opportunity is obtained by controlling the antenna size electrically, or adding some mechanical structure to the antenna like resistors, varactor diodes, PIN diodes, RF-MEMS switches, and tunable electromagnetic materials (e.g., graphene, liquid crystal (LC) [4], and ferroelectric film) [5, 6]. Furthermore, those antennas are low profile, have low power losses, low cost, and easily manufactured with standard PCB techniques.

Ideally, a circularly polarized reconfigurable patch antenna with an inhomogeneous substrate has been investigated in [7], operating at 2.4 GHz for RFID application. A reconfigurable feeding antenna is presented in [1], sixteen PIN diodes are used to control the direction of the radiation pattern. In [8], omnidirectional and directional operational modes are realized, and a beam scanning of a complete angle of 360◦ is achieved by using arc-shaped dipoles and a circular patch used as a reflector.

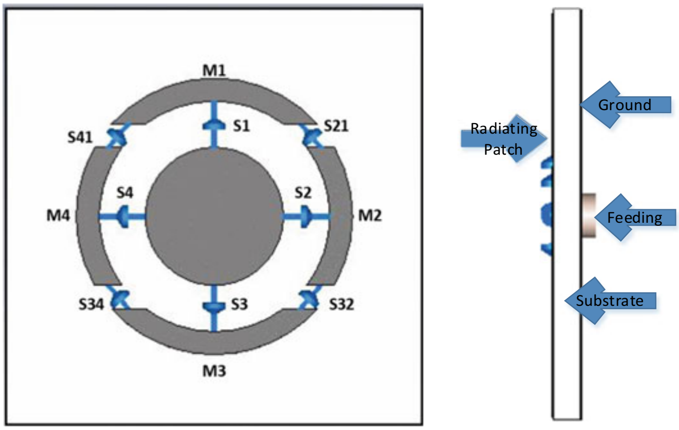
This paper reports a versatile, electrically small, and novel frequency reconfigurable circular microstrip antenna structure for the multiple-band operation of (5G) 5.2 GHz ISM bands, 7.5, 8.7, 10.2, 10.7, and 14 GHz. Tunable PIN diodes have been proposed for integrating circular and identical microstrips. Each diode can be separately switched ON/OFF. While the switch is in the ON state, the current distribution on the structure of the antenna is extremely different than if the switch is in the OFF state.

2 Reconfigurable Antenna Design

Antenna design requires a designer sense and high computational potential to define the design elements of an antenna that satisfies a reasonable performance for a required operating frequency range. The studied structure presents a novel and electrically small reconfigurable circular microstrip antenna as shown in Fig. 1. The antenna is often electrically small if Eq. (1) is fulfilled [9, 10].

$$\frac{a}{\lambda} = \frac{1}{2\pi} \tag{1}$$

where a is the radius of the circular part set to 0.3 cm and λ is the wavelength set to 5.6 cm for a reference frequency of 5.2 GHz. The proposed antenna has been developed to manipulate various antenna elements effectively and to increase performance by controlling identical microstrips with PIN diodes. It has three layers



(a) (b)

Fig. 1 Reconfigurable antenna. **a** Top view and **b** side view

(radiating, substrate, and reflector) with feeding at the center of the circular patch, integrated by four switches (S_1, S_2, S_3, S_4) to four identical microstrip elements (M_1, M_2, M_3, M_4). The radiating conductors replaced on the top of the substrate consist of a circular patch with different four identical microstrip elements integrated also by four switches (S_{21}, S_{32}, S_{34}, S_{41}) to each other. They are modeled on Rogers RT5880LZ substrate with dimensions of 1.8×1.8 cm^2, a thickness of 0.3175 cm and a permittivity ε_r of 1.96. The reflector is printed on the backside of the substrate. The unfilled spaces between the circular and identical microstrip elements set to 0.2 cm, in addition to unfilled spaces between identical microstrip elements set to 0.132 cm include four switches. The geometry of circular microstrip has a radius a of 0.3 cm ($0.053\lambda_0$) and a height h of 0.006604 cm ($0.0011\lambda_0$). The resonant frequency f_r of a circular microstrip antenna is given by [11]:

$$f_r = \frac{K_{mn} c}{2\pi a_e \sqrt{\varepsilon_r}} \tag{2}$$

where a_e is an effective radius for the circular microstrip, c is velocity of light in free space and $K_{mn} = m$ th zero of the derivative of Bessel function at order n. The fundamental mode is TM_{11} for K set to 1.84118. a_e is given by [12]:

$$a_e = a\left\{1 + \left(\frac{2h}{\pi a \varepsilon_r}\right)\left[\ln\left(\frac{\pi a}{2h}\right) + 1.7726\right]\right\}^{1/2} \tag{3}$$

Switches positioned on the structure are RF-PIN diode switches (SW1AD-33) with frequency range of 0.3–18 GHz and switching time 100 ns [13]. Switching states is ON state (RL serial-forward bias) while the resistor R has a value of 5

Ohms and 1000 Ohms for OFF state (RLC parallel-reverse bias) [6]. The action of the biasing process is based on supplying the current to the integrated switching elements used to configure the antenna to achieve new frequency operations. The value of inductance L (0.1 nH) and capacitance C (0.3 pF) has been avoided in the simulation. The current distributions on identical microstrips are by integrated resistances as shown on Fig. 1.

3 Results and Discussion

The proposed reconfigurable circular antenna includes eight switches and two states of switching either ON or OFF. As the state of the antenna is changed with switches, so the current distribution on the connected elements is changed. In these cases, the changes in antenna parameters by switches can be used to change the current distribution states on the conductors of the antenna, which then result in operating frequency changes. The antenna's performance has been reported by the S-parameter curves as presented in the following figures and discussion.

At first, Fig. 2 shows the result of the simulation for the circular microstrip only, while all switches is in OFF state. Operating frequency is 5.2 GHz with a bandwidth of 3 GHz, at a target of return loss ≤ -10 dB. The reported operating frequency can be used for the fifth generation (5G) ISM 5.2 GHz in wideband.

Secondly, Fig. 3 shows a new result that depends on switching S_1 to ON state, which allows the current transition to an identical microstrip of M_1, while others (M_2, M_3 and M_4) are independents. Operating frequency is 7.5 GHz with a bandwidth of 2.44 GHz, at a target of return loss ≤ -10 dB.

Thirdly, S_1, S_2, S_3, and S_4 are four evenly accesses for current flowing in state ON through M_1, M_2, M_3, and M_4. Accordingly, the proposed antenna resonates at high frequency of 14 GHz with a bandwidth of 2.3 GHz as shown in Fig. 4.

Fourthly, in the case of the current flowing only through S_1, and S_3 to the identical microstrips of M_1 and M_3, the operating frequency is 10.2 GHz with a bandwidth of 3 GHz, and if only S_2 and S_4 are in ON state, the operating frequency is 10.7 GHz

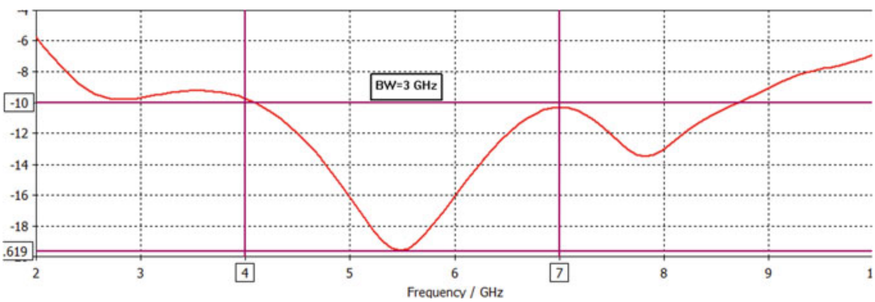


Fig. 2 S-parameter, all switches is in OFF state

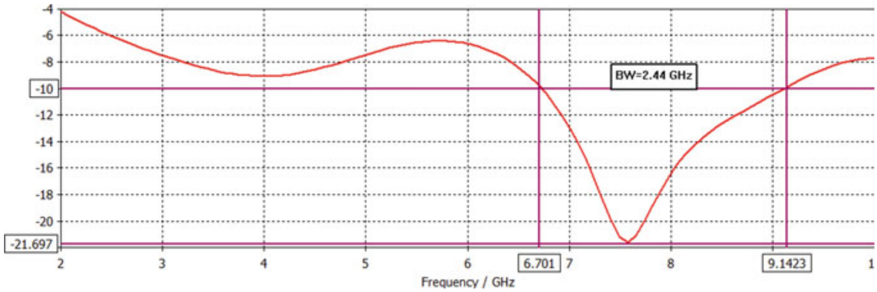


Fig. 3 S-parameters, S_1 is only switched ON others are OFF

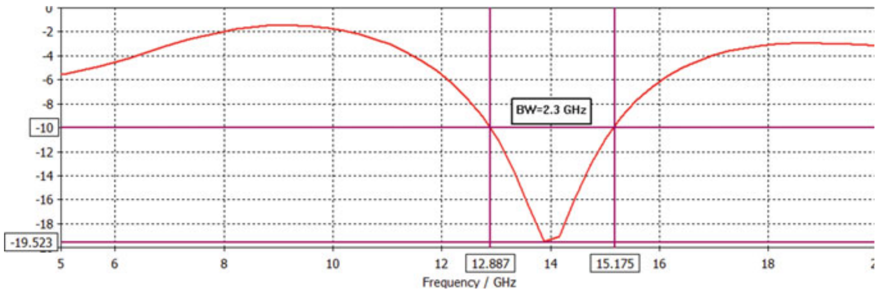


Fig. 4 S-parameter, S_1, S_2, S_3, and S_4 are switched ON others OFF

with a bandwidth of 2.8 GHz as shown in Fig. 5. It is noticed that although the symmetricity of the form and the current flowing are identical once by S_1 and S_3 and once by S_2 and S_4, the results are not equal. The inequality of the results is due to the difference in the area of each of M_1 and M_3 with that of M_2 and M_4.

Fifthly, in the case of S_1,S_3, S_{21},S_{32}, S_{34}, and S_{41} are only switched ON, the optimum operating frequency is 8.78 GHz with a bandwidth of 3.2 GHz. The current transition flows through M_1 and M_3 to M_2 and M_4. If only S_2, S_4, S_{21},S_{32}, S_{34} and

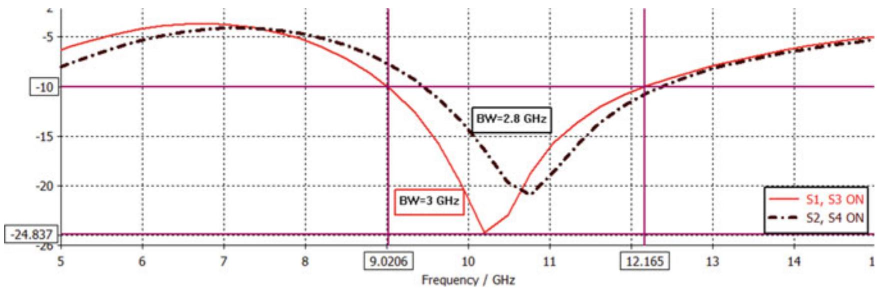


Fig. 5 S-parameters, (1) S_1, and S_3 are switched ON others OFF, (2) S_2, and S_4 are switched ON others OFF

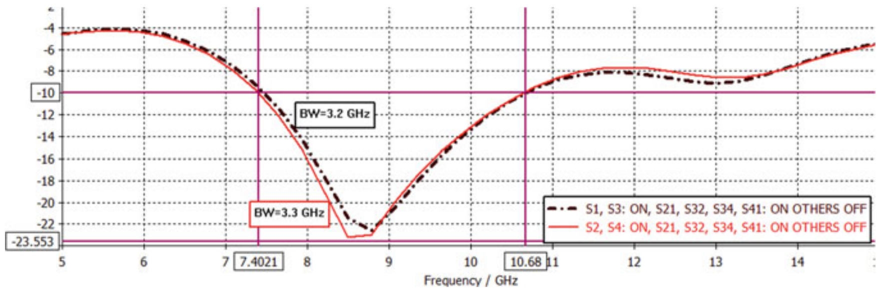


Fig. 6 S-parameters, (1) S_1, S_3 S_{21}, S_{32} S_{34}, and S_{41} are switched ON others OFF, (2) S_2, S_4 S_{21}, S_{32} S_{34}, and S_{41} are switched ON others OFF

S_{41} are switched ON, the optimum operating frequency is 8.5 GHz with a bandwidth of 3.3 GHz as shown in Fig. 6. The current transition flows through M_2 and M_4 to M_1 and M_3. The slight difference in results is due to the difference in the area of each of M_1 and M_3 with that of M_2 and M_4.

The figures above show the performance of the antenna during simulation. Switching the structure parameters results in novel operating frequencies without changing antenna characteristics physically. The operating bands of 5.2 GHz ISM bands, 7.5, 8.7, 10.2, 10.7, and 14 GHz are considered for many wireless communication applications, such as MIMO systems, cognitive radio system, satellite, biomedical, and industrial applications. This confirms the advantage of the antenna as a multiple band frequency reconfigurable antenna with a less complexity due to using eight PIN diodes as switches. This antenna is different from other conventional antennas since its control part for reconfiguring is found in the antenna itself.

4 Conclusion

A frequency reconfigurable circular microstrip antenna is modeled and simulated for wireless communication and mobile devices. Eight PIN diodes is integrated between the circular and identical microstrips to obtain new operating frequencies. The PIN diode works in two switching positions, and therefore the proposed antenna works in the two positions as well. In each position, a new operating frequency is generated. The proposed reconfigurable antenna provides a multiple operating range between 5 and 15 GHz. Since the proposed reconfigurable antenna has five different structures, it produces six different operating frequencies. The dimensions of the proposed antenna are actually small, so further research can be done on the antenna in order to be suitable for the future 5G and next-generation market. It is prospected that the necessity for reconfigurable antennas can also encourage development in several wireless communication areas.

References

1. Bernhard JT (2007) Reconfigurable antennas. Champaign, Morgan & Claypool Publishers. https://doi.org/10.1002/0471654507.eme514
2. Costantine J (2009) Design, optimization and analysis of reconfigurable antennas. Ph.D. Albuquerque, New Mexico
3. Aoad A, Aydın Z (2020) New modeling of reconfigurable microstrip antenna using hybrid structure of simulation driven and knowledge based artificial neural networks. Pamukkale Univ J Eng Sci 5:935–943. https://doi.org/10.5505/pajes.2020.67809
4. Costantine J, Tawk Y, Christodoulou CG (2013) Design of reconfigurable antennas using graph models. Springer, Cham. https://doi.org/10.1007/978-3-031-01540-3
5. Aoad A, Aydın, Z, Korkmaz E (2014) Design of a tri band 5-fingers shaped microstrip patch antenna with an adjustable resistor. In: IEEE antenna measurements & applications (CAMA). Antibes Juan-Les-Pins. doi:https://doi.org/10.1109/CAMA.2014.7003444
6. Singh A, Dubey R, Jatav R, Meshram MK (2022) Electronically reconfigurable microstrip antenna with steerable beams. Int J Electron Commun 149:1–8. https://doi.org/10.1016/j.aeue.2022.154179
7. Chen Z, Li HZ, Wong H, Zhang X, Yuan T (2021) A circularly-polarized-reconfigurable patch antenna with liquid dielectric. IEEE Open J Antennas Propag 2:396–401. https://doi.org/10.1109/OJAP.2021.3064996
8. Miao X, Wan W, Duan Z, Geyi W (2019) Design of dual-mode arc-shaped dipole arrays for indoor base-station applications. IEEE Antennas Wirel Propag Lett 18:752–756. https://doi.org/10.1109/LAWP.2019.2901967
9. Huang Y, Boyle K (2008) Antennas: from theory to practice. Wiley
10. Miron DB (2006) Small antenna design. Elsevier Inc. https://doi.org/10.1016/B978-0-7506-7861-2.X5000-4
11. Kumar G, Ray K (2002) Broadband microstrip antennas. Artech
12. Garg R, Bhartia P, Bahl I, Ittipiboon A (2001) Microstrip antenna design hand book. Artech House
13. Pulsar Microwave Corporation (2002). Available https://www.pulsarmicrowave.com/product/switch/SW1AD-33

Online Protection for Children Using a Developed Parental Monitoring Tool

Martin Stoev and Dipti K. Sarmah

Abstract Nowadays, children are more comfortable using online tools for their education. This is also enhanced during the COVID-19 period when one must do online activities such as calls, studies, and meetings. Children who use the Internet regularly, especially in the age group between 6 and 14, may experience Internet risk. While on the Internet, children can fall victim to hateful, age-inappropriate content, cyberbullying, phishing, etc. There are also risks of privacy violations by the websites which can be done through cookies or user account browsing features. Further, excessive use of the Internet may negatively impact a developing child's physique, cause sleeping problems, and potentially lead to addiction. To protect their children, some parents use Android applications such as Google Family Link (Free), Kids Place Parental Control (free and paid), Norton Family (paid), Qustodio (paid), and FamiSafe (paid). These applications allow the parents to restrict and monitor the child's behavior. However, many features are not implemented in free applications such as monitoring calls, messages, social media, etc. This research aims to analyze the Google Play Store reviews (positive and negative) of popular and mentioned Android applications and their various features. Based on the research, a free alternative in the form of an application was developed. It has two components—an Android application for the child's device and a web interface for the parent. This tool allows parents to monitor calls, contacts, and SMS.

Keywords Parental monitoring applications (free and paid) · Children online protection · Android application · Google play store reviews

M. Stoev (✉) · D. K. Sarmah
Services and Cyber Security Department, University of Twente, Enschede, The Netherlands
e-mail: m.stoev@student.utwente.nl

D. K. Sarmah
e-mail: d.k.sarmah@utwente.nl

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_17

1 Introduction

In modern society, many young people participate in online activities such as studying and relaxing. Children tend to invest their time in games, videos, and chatting with their friends [1]. Parents are also concerned about the behavior of their children when they spend most of their time online. They want to give them a safe online environment by ensuring their children's online privacy and keeping them away from malicious content, and inappropriate content [2]. Children can also suffer from having their personal information and photo shared on the Internet [3]. Such leaks may lead to the child experiencing cyberbullying, [4] where a parent's involvement is crucial. Data collection may violate their privacy, and lead to bullying [3]. There have been published websites on children's protection laws that protect the privacy of children. Some of the most popular is the Children's Online Privacy Protection Act (COPPA) [5] and the General Data Protection Regulation (GDPR) [6]. However, they are not fully compliant [7]. Therefore, there is a need to monitor the child's behavior to which many parents take a restrictive approach [1]. They prohibit websites that may be harmful to the child and limit the amount of time spent on the phone. This is done to avoid developing sleeping [1] and physical [8] problems, addiction, antisocial behavior [8, 9], and being exposed to violence [8]. Many parents take plenty of factors into consideration when choosing applications for their children. Privacy and parental permission required for tasks, such as shopping and age-appropriate content are seen as the most crucial factors for parents [1].

A lot of parents express the need to view their children's chat history and want to know for what purposes their children's personal information is being used [3]. Unfortunately, a lot of parents are not able to enable parental control on their children's mobile phones. Also, many of them feel societal pressure to allow their children to make social media accounts, exposing them to more potential risks [3]. Many applications [10–15] aim to solve this problem. The parental control applications that are most downloaded and reviewed in the Google Play Store are Google Family Link [10], Kids Place Parental Control [11], Norton Family [12], Qustodio [13], and FamiSafe [15]. All of these parental monitoring applications are capable of monitoring and controlling screen time (see Tables 1 and 2). Google Family Link and Kids Place Parental Control require parental permissions to install applications [10, 11]. FamiSafe [15] is seen as the best parental monitoring application [16] because of the amount of features, it has (see Table 2).

Unfortunately, there are many limitations in popular parental monitoring applications [10–15]. While Google Family Link [10] is free and with a limited set of features, others provide a wider range of features at a high price. The web filters Google Family Link provides are very weak, letting many age-inappropriate content pass through them [17]. Parents can restrict and monitor the child only when they are not older than 13 [17]. Once the child becomes 13 years old, they are able to avoid being monitored by their parents. In addition to that, some reviews on the Google Play Store express frustration over the lack of functionality and occasional malfunctions [10] (See Table 1). Kids Place Parental Control shares these shortcomings in the free

Table 1 Price, features, and reviews of parental monitoring applications

Price and features	Praised aspects	Disliked aspects
Google family link [10] Free [10] • Completely free • Good design • Location tracker • Restrict and monitor screen time • Usage schedule • Web filters (no categories)	**Positive reviews: 77%** • **47%** generally positive • **14%** restrictive features • **9%** monitoring features • **3%** good interface • **2%** web filters • **2%** location tracker	**Negative reviews: 22%** • **11%** of bugs occurred • **6%** generally negative • **2%** bugs occurred with the location tracker • **2%** disliked the interface • **1%** small set of features
Kids place [11] $4.99/month [18] • Monitor application usage(free) • Child lock and YouTube Safe search(free) • Web filtering(paid) • Location tracker (paid) • Web site access reports(paid) • Control application usage(paid)	**Positive reviews: 53%** • **33%** generally positive • **15%** restrictive features • **3%** good design • **2%** web filters	**Negative reviews: 47%** • **16%** of bugs occurred • **9%** disliked the design • **6%** generally negative • **6%** of bugs occurred with restrictive features • **5%** high price • **5%** easy to bypass
Norton family [12] €3999/year [23] • Location features • Monitor and control • Screen time per application • Web filters • Location tracker • Track browser history	**Positive reviews: 30%** • **21%** generally positive • **3%** restrictive features • **2%** monitoring features • **2%** web filters • **1%** location tracker • **1%** good design	**Negative reviews: 70%** • **22%** of bugs occurred • **20%** generally negative • **9%** of bugs occurred with restrictive features • **8%** disliked the design • **4%** hard to configure • **4%** easy to bypass • **3%** bugs occurred with the location feature

version, however, has them implemented in the premium version [18] (see Table 1). Norton Family is a premium application (see Table 1) with a poorly designed user interface design and occasionally malfunctions (see Table 1). Few of these applications can monitor chats such as WhatsApp [19]. FamiSafe [15], however, provides these features alongside many others at the cost of 60 Euros yearly (see Table 2).

The research aims to evaluate said parent monitoring applications and recognize important and missing aspects of parental monitoring applications. Based on the results, an alternative application is developed, addressing the desired features with a user-friendly interface. This research is divided into these research questions (RQ):

RQ1. What aspects do parents value in parental monitoring applications?
RQ2. What aspects of parental monitoring applications need to be improved or implemented?
RQ3. Can an alternative be implemented that improves upon existing parental monitoring applications?

Table 2 Price, features, and reviews of parental monitoring applications

Price and features	Praised aspects	Disliked aspects
Qustodio [13] **€42.95/month** [24] • App and content filtering • Activity monitoring • Setting usage limits • Calls/SMS monitoring • Location tracker • Warnings for threatening activities • Chat monitor (WhatsApp, messenger)	**Positive reviews: 46%** • **32%** were generally positive • **8%** restrictive features • **3%** good design • **2%** monitoring features • **1%** location tracker	**Negative reviews: 49%** • **11%** of bugs occurred with the restrictive features • **11%** high price • **8%** generally negative • **7%** is bypassed easily • **4%** of bugs occurred • **4%** hard to configure • **4%** of bugs occurred with the monitoring features
FamiSafe [15] **$10.99/month** [25] • Location tracker and gallery monitoring • Activity monitoring and web filters • App blocker and screen time control • Monitor social media for suspicious texts • Browser history monitoring • YouTube monitoring and restricting	**Positive reviews: 84%** • **24%** monitoring features • **22%** location tracker • **15%** generally positive • **14%** restrictive features • **5%** good design • **2%** YouTube restrictions • **2%** web filters	**Negative reviews: 16%** • **7%** high price • **4%** generally negative • **2%** of bugs occurred • **3%** hard to configure

2 Related Work

During an analysis of reviews of some parental monitor applications, Alelyani et al. [20] found that parents are dissatisfied with the cost, performance, and ease of use of the parental monitoring applications they use. On the other hand, adolescents are dissatisfied with the effects of their parental monitoring application on their autonomy. McNally et al. [21] also discussed the same concerns. However, this research could not identify what features were missing in any parental monitoring applications.

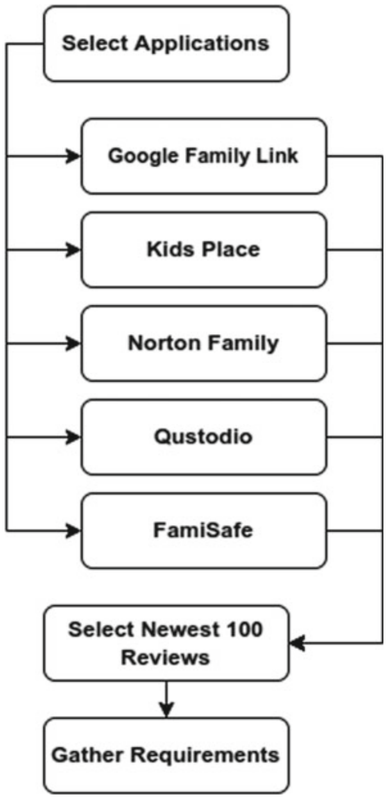
In addition, research has been conducted on how certain parental control applications achieve privacy and security [22]. The authors considered parental control applications Norton Family. This research has discovered that Norton Family, upon visiting the website, does not encrypt the user data, exposing the user to a major privacy threat. In relation to the development of the application, Warner et al. [26] plan the stages and provide facilities for the development of Android monitoring applications. Instructions such as monitoring calls and browser history are considered. None of the mentioned parental monitoring applications directly protect children's privacy. Parental Online Consent for Kid's Electronic Transactions (POCKET) [27] addresses this directly. It possesses several significant aspects being able to automate parental consent, managing priorities for collected data, viewing the data that is being stored, and verifying the privacy practices of a given website. Instructions from the development of POCKET were considered. SafeChat [28] takes a preventive approach to cyberbullying by censoring harmful words and providing a safe environment for the child. The research [28] gives instructions for achieving censoring and security, which were taken into consideration.

The next section discusses how the research questions are answered. This will be done by analyzing popular parental monitoring applications and the reviews parents have given them.

3 Methodology

This section is dedicated to the methodology of research questions, also shown in Fig. 1. The existing applications and their features are also analyzed in Sect. 3.1.

Fig. 1 Methodology visualization



Step 1—RQ1, RQ2

These questions are answered by categorizing the application's reviews on Google Play Store as mentioned in Sect. 1 [10–15].

For every application, the latest 100 reviews are picked. This is done to determine the consensus on the latest features of each application.

Step 2—RQ3

Using the answers to the first two research questions, an application is created, aiming to solve some of the lacking aspects. These aspects were determined by negative reviews and the reviews that request missing features.

3.1 Evaluation of Existing Solutions

The parental monitoring applications [10–15] with the highest number of downloads and comments on the Google Play Store were chosen for analysis of their reviews and features (see Tables 1 and 2).

Requested Features from Popular Applications' Reviews

- Requested from four reviews to have a more efficient aggregation of chat messages in Qustodio.
- Requested from one review to have pictures shown along with chat messages in Qustodio
- Requested from one review to have monitoring of calls in Google Family Link

The reviews of all these applications will be discussed more in-depth in the next section.

4 Results

In this section, the parent's reviews will be analyzed to answer research question 1 and research question 2. Using the analysis, a developed free alternative will be discussed in Sect. 5 that aims to help parents with monitoring their children.

4.1 Research Question 1

Considering the analysis of the 500 reviews in Tables 1 and 2, the value of a feature is determined by the frequency with which it is mentioned by different reviews.

Ranking these features by this value gives the following list (features can also be seen in Tables 1 and 2):

1. **Restrictive**—limit the time a given application is used (**16.0%**—54 positives, 26 negatives).
2. **Monitoring**—see how much time is spent per application (**8.2%**—37 positives, 4 negatives).
3. **Locating**—receive alerts when children leave a set perimeter and track the visited locations (**6.2%**—26 positives, 5 negatives).
4. **Web Filters**—limit the usage of websites containing a specified category such as pornography or gambling (**6.0%**—8 positives, 0 negative).
5. **Chat Monitoring**—keep track of what texts are being exchanged (**1.0%**—5 feature requests)
6. **Blocking YouTube channels** (**0.4%**—2 positive, 0 negative).
7. **Calls/SMS Monitoring** (**0.2%**—1 feature request).

4.2 Research Question 2

Considering the analysis from Tables 1 and 2, the list below lists the features that need to be improved according to parents. The items on this list are ranked by having the feature's number of negative reviews divided by the frequency with which it is mentioned. The last two are requests from parents.

1. **Restrictive Features** (32.5%)
2. **Location Features** (19.2%)
3. **Monitoring Activity** (9.8%)
4. **Monitoring Messages** (1.0%—5 feature requests)
5. **Calls/SMS Monitoring** (0.2%—1 feature request)

Additional requirements have been gathered based on other negative reviews:

1. Consistent functionality (55%)
2. Affordable price (23%)
3. Good design (19%)
4. Hard to bypass (16%)
5. East to configure (11%)

5 Developed Solution

Based on the answers to research question 2, desired features that are not present in any free applications are blocking contacts and monitoring SMS and calls. This section discusses the development of an application for the child's Android phone and a web application for parents to monitor contacts, SMS, and calls. The Android application requires Android version 21 or above.

5.1 The Child's Android Application

The application gathers data about SMS and calls and uploads the information to a Firebase Realtime database which is also compliant with GDPR [29]. The Firebase Realtime database stores the data per mobile phone and allows parents who view the web application to receive updates in real time. The application was developed using Android Studio Bumblebee 2021.1.1 and Java 17. Due to the main idea of the application being to collect data, no user interface was developed. When installing the application, the following permission needs to be granted: calls logs, contacts, SMS, and phone. Blocking contacts was an intended feature that was not possible to be implemented. This is because all the possible solutions do not work on the current version of Android Studio. During the first installation, parents must authenticate themselves with their Google accounts. This account will also be used for the web application.

5.2 The Parent's Web Application

Using their Google account, parents can log in to the web application and monitor their children's SMS and calls. Logging in is done using a pop-up (the web application requires the user to allow pop-ups to be shown). It was developed using Svelte [30], a JavaScript framework, and Tailwind [31], a CSS library. The rules set in the Firebase Realtime Database allow accountants to be able to access and modify only their information, protecting users from outside risks [32]. The data is stored in the plain text due to the short development period. The sidebar of the website has tabs with all the website's pages. The first is the call tab (see Fig. 2) which provides an analysis of the calls for the current month and a column with all the calls. Every call displays the contact, date, and duration of the call.

The distinct types of calls are also shown: REJECTED, INCOMING, OUTGOING, MISSED. The list can be filtered by name and phone number. Each call has a button with a lock icon next to it. That button is supposed to block the contact, but this functionality cannot be implemented anymore on Android. The second tab is contacts/SMS (see Fig. 3). It has a column for SMS and contacts and can also be filtered. The block icon can also be seen next to each contact. The list of SMS possesses information about who the sender is, when it is sent, the body of the SMS, and whether it was RECEIVED or SENT.

5.3 Discussion

The implemented alternative does not possess plenty of premium features, it makes up for it by being easy to configure, well designed, and free. Parents who are interested in

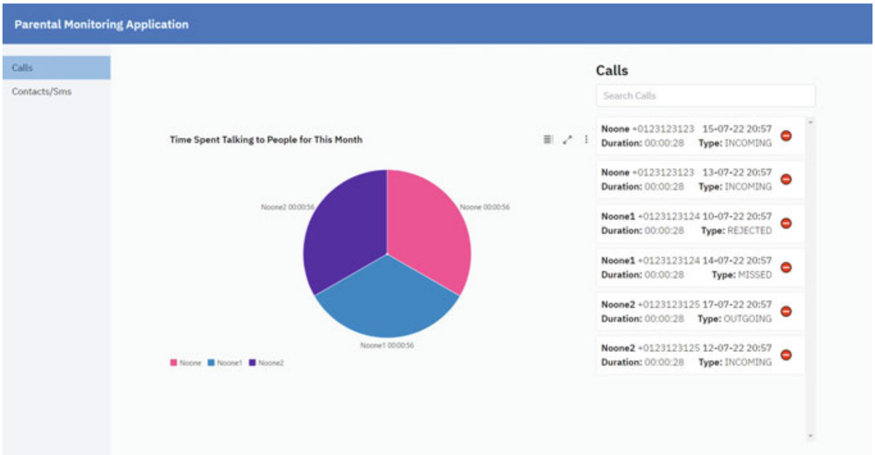


Fig. 2 Web application calls page

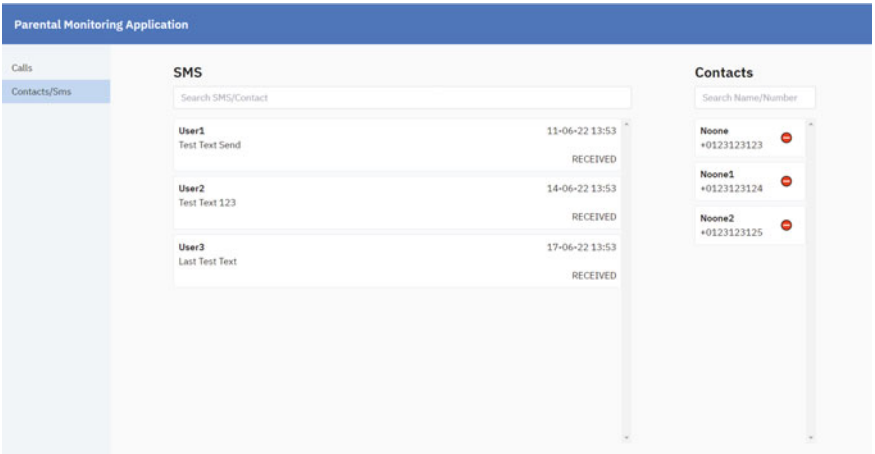


Fig. 3 Web application contacts/SMS page

personally testing the application can find the *application's apk* (https://drive.google.com/file/d/1zIyCMq8BcwElUdcnNRtt7SNQUOcdz_lQ/view?usp=sharing) and the *web application* (https://candid-frangipane-13f547.netlify.app/) are online. Permissions for the applications can be given upon application launch or in the settings. The code for the *Android application* (https://github.com/MartinStoev00/ParentAndroid) and *Svelte web application* (https://github.com/MartinStoev00/ParentSvelte) are both available online.

6 Conclusion and Future Work

This research found missing features in popular parental monitoring applications (Google Family Link, Kids Place Parental Controls, Norton Family Parental Control, Qustodio, FamiSafe) such as monitoring calls, SMS, messages, social media monitoring, and more. Based on these limitations, a call, SMS, and contacts monitoring Android application were developed. The developed tool lacks a wide range of features; however, parents can still use it to protect their children. For future work, the focus is on implementing the blocking of contacts and other premium features.

References

1. Brito R, Dias P (2020) Which apps are good for my children? How the parents of young children select apps. In J Child–Comput Interact 26:100188. Available https://www.sciencedirect.com/science/article/pii/S2212868920300180
2. Dias P, Brito R (2021) Criteria for selecting apps: Debating the perceptions of young children, parents and industry stakeholders. Comput Educ 165:104134. Available https://www.sciencedirect.com/science/article/pii/S0360131521000117
3. Manotipya P, Ghazinour K (2020) Children's online privacy from parents' perspective. Procedia Comput Sci 177:178–185. In: The 11th international conference on emerging ubiquitous systems and pervasive networks (EUSPN 2020)/The 10th international conference on current and future trends of information and communication technologies in healthcare (ICTH 2020)/affiliated workshops
4. Elgar FJ, Napoletano A, Saul G, Dirks MA, Craig W, Poteat VP, Holt M, Koenig BW (2014) Cyberbullying victimization and mental health in adolescents and the moderating role of family dinners. JAMA Pediatr 168(11):1015–1022
5. Children's online privacy protection act, public law No. 105–277, 112 stat. 2681-728. 1998. Available https://www.govinfo.gov/content/pkg/FR-2010-04-05/html/2010-7549.htm
6. General Data Protection Regulation (GDPR), Article 8—Conditions applicable to child's consent in relation to information society services. Available https://gdpr-info.eu/art-8-gdpr/
7. Cai X, Gantz W, Schwartz N, Wang X (2003) Children's website adherence to the FTC's online privacy protection rule. J Appl Commun Res 31(4):346–362. Available https://doi.org/10.1080/1369681032000132591
8. Ankaya SC, Odabaşı (2009) Parental controls on children's computer and Internet use. Procedia Soc Behav Sci 1(1):1105–1109. In: World conference on educational sciences: new trends and issues in educational sciences. Available https://www.sciencedirect.com/science/article/pii/S187704280900202X
9. Cho K-S, Lee J-M (2017) Influence of smartphone addiction proneness of young children on problematic behaviors and emotional intelligence: Mediating self-assessment effects of parents using smartphones. Comput Human Behav 66:303–311. Available https://www.sciencedirect.com/science/article/pii/S0747563216306987
10. Google family link—Apps on Google Play. Available https://play.google.com/store/apps/details?id=com.google. Last accessed 7 Sept 2022
11. Kids place parental controls—Apps on Google Play. Available https://play.google.com/store/apps/details?id=com.kiddoware.kidsplace&gl=US. Last accessed 7 Sept 2022
12. Norton family parental control—Apps on Google Play. Available https://play.google.com/store/apps/details?id=com.symantec.familysafety. Last accessed 7 Sept 2022
13. Qustodio. Available https://www.qustodio.com/en/. Last accessed 7 Sept 2022

14. Parental control—Screen time location tracker. Available https://play.google.com/store/apps/details?id=com. Last accessed 7 Sept 2022
15. FamiSafe: parental control app. https://play.google.com/store/apps/details?id=com.wondershare.famisafe&gl=US. Last accessed 7 Sept 2022
16. Family choice awards. https://www.familychoiceawards.com/family-choice-awards-winners/wondershare-famisafe/. Last accessed 7 Sept 2022
17. Google family link—FAQ. https://families.google.com/familylink/faq/. Last accessed 7 Sept 2022
18. Kiddoware pricing plans—Parental control apps for Android. https://kiddoware.com/pricing-plans-kids-place-safety/. Last accessed 7 Sept 2022
19. WhatsApp messenger. https://play.google.com/store/apps/details?id=com.whatsapp. Last accessed 7 Sept 2022
20. Alelyani T, Ghosh AK, Moralez L, Guha S, Wisniewski P (2019) Examining parent versus child reviews of parental control apps on Google Play, pp 3–21
21. McNally B, Kumar P, Hordatt C, Mauriello ML, Naik S, Norooz L, Shorter, Golub E, Druin A (2018) Co-Designing mobile online safety applications with children, pp 1–9. https://doi.org/10.1145/3173574.3174097
22. Feal Fajardo A (2017) Study on privacy of parental control mobile applications. Ph.D. dissertation. ETSI Informatica
23. Norton family pricing. http://nl.norton.com/products/norton-family?inid=familycom_subscribe_home. Last accessed 7 Sept 2022
24. Qustodio price. Available https://www.qustodio.com/en/premium/. Last accessed 7 Sept 2022
25. FamiSafe: parental control app pricing. Available https://famisafe.wondershare.com/store/family.html. Last accessed 7 Sept 2022
26. Warner T, Meadows C, Wahjudi P (2012) Analysis, recognition, monitoring, and reporting tool (ARMR). J Manage Eng Integr 20
27. Bélanger F, Crossler RE, Hiller JS, Park JM, Hsiao MS (2013) POCKET: a tool for protecting children's privacy online. Decis Support Syst 54(2):1161–1173, 2013. Available https://www.sciencedirect.com/science/article/pii/S0167923612003429
28. Fahrnberger G, Nayak D, Martha VS, Ramaswamy S (2014) SafeChat: a tool to shield children's communication from explicit messages, pp 80–86
29. Privacy and security in firebase. Available https://firebase.google.com/support/privacy
30. Svelte. Available https://svelte.dev/. Last accessed 7 Sept 2022
31. TailwindCSS. Available https://tailwindcss.com/. Last accessed 7 Sept 2022
32. Understand firebase realtime database rules. Available https://firebase.google.com/docs/database/security. Last accessed 7 Sept 2022

12 bit 1 ps Resolution Time-to-Digital Converter for LSI Test System

Daisuke Iimori, Takayuki Nakatani, Shogo Katayama, Misaki Takagi, Yujie Zhao, Anna Kuwana, Kentaroh Katoh, Kazumi Hatayama, Haruo Kobayashi, Keno Sato, Takashi Ishida, Toshiyuki Okamoto, and Tamotsu Ichikawa

Abstract This paper describes a 12 bit, 1 ps resolution, 5 ns full-scale time-to-digital converter (TDC) for LSI test system application. The TDC is realized with discrete electronic components on a board for low cost, which is suitable for LSI test system application and expected to use as built-in self-test circuit (BIST). In the TDC, the upper 9 bits are obtained by successive approximation register (SAR) configuration using 9-bit programmable variable delay elements, while the lower 3

D. Iimori (✉) · T. Nakatani · S. Katayama · M. Takagi · Y. Zhao · A. Kuwana · K. Katoh · K. Hatayama · H. Kobayashi
Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, 1-5-1 Tenjin-Cho Kiryu Gunma, Maebashi 376-8515, Japan
e-mail: t170d011@gunma-u.ac.jp

S. Katayama
e-mail: t15304906@gunma-u.ac.jp

M. Takagi
e-mail: t221d046@gunma-u.ac.jp

Y. Zhao
e-mail: t181d041@gunma-u.ac.jp

A. Kuwana
e-mail: kuwana.anna@gunma-u.ac.jp

H. Kobayashi
e-mail: ikoba@gunma-u.ac.jp

K. Sato · T. Ishida · T. Okamoto · T. Ichikawa
ROHM Semiconductor, 2-4-8 Shin-Yokohama, Kouhoku-Ku, Yokohama 222-8575, Japan
e-mail: keno.sato@dsn.rohm.co.jp

T. Ishida
e-mail: takashi.ishida@lsi.rohm.co.jp

T. Okamoto
e-mail: toshiyuki.okamoto@lsi.rohm.co.jp

T. Ichikawa
e-mail: tamotsu.ichikawa@lsi.rohm.co.jp

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_18

bits are by injecting jitter at the TDC input, measuring 100 times and estimating the most probable digital value with the statistical processing. The prototype TDC performance is evaluated with experiments.

Keywords Time measurement · TDC · SAR · Vernier TDC · Jitter injection · Cumulative distribution function · LSI test system · BIST

1 Introduction

A time-to-digital converter (TDC) is used to measure the time difference between the edges of two timing signals as a digital output (Fig. 1). A TDC with 1 ps resolution can be achieved in advanced LSI process, but its development cost is extremely high and development time is long; it can be used such as in consumer electronics products, where large volumes are shipped. However, it is not for some LSI test systems, where such large volumes are not used [1]. In this paper, we show a TDC using discrete electronic components with comparable performance to the one on advanced process full custom IC. Our design techniques, implementation and measurement results are shown.

2 Proposed TDC Architecture

2.1 SAR TDC for Upper Bits

Before going into the SAR TDC, a description of the SAR TDC is given. In the SAR TDC, the sampled analog signal and the output of DAC are sequentially compared starting from MSB to LSB so that they match.

- Sample and hold the analog input voltage signal.

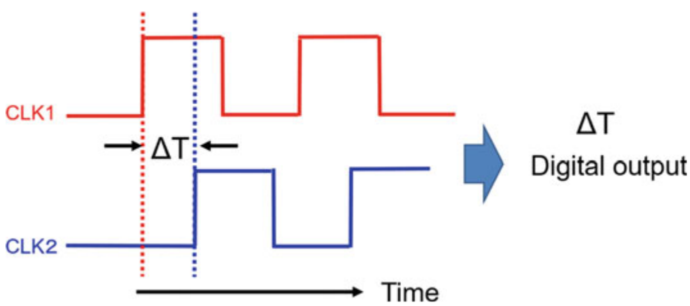


Fig. 1 Time difference measurement by TDC

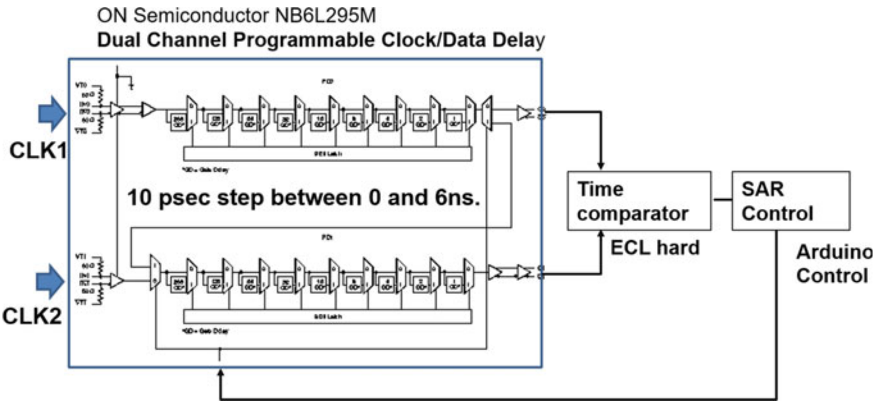


Fig. 2 Prototype SAR TDC configuration to obtain the upper 9 bits

- Set “1” to MSB in the SAR logic.
- The digital value from the SAR control logic is converted to an analog value by the internal DAC
- Compare the sampled voltage with the DAC output voltage.
- In case, the sampled voltage is larger than DAC output voltage, set MSB to “1”.
- Otherwise, set MSB to “0”.

The above operation is repeated from MSB to LSB.

The SAR TDC uses this principle of successive comparison to measure the time difference between two clocks (Fig. 1) [2–6].

Now the circuit operation of the prototype SAR TDC is explained (Fig. 2). The two input clock signals CLK1, CLK2 are provided to dual 9-bit programmable variable delay device; one is set as the reference (delay = 0) and the other is set to variable delay with 9-bit resolution ($n = 0$–511) by Arduino control. Both outputs go to the comparator (D Flip-Flop) and its output is provided to the SAR control logic. Based on this, the SAR control logic outputs a multiplexer selection signal in the dual 9-bit programmable variable delay element based on the binary search principle. The 9-bit digital output can be obtained by repeating these operations [7].

2.2 Vernier TDC for Lower Bits

In this section, the Vernier TDC circuit and operation to obtain fine time resolution as lower 3 bits is explained (Fig. 3) [8–10].

1. Search for a value as the same delay value as sig. clock by varying the delay value in the SAR sequence for ref clock. The delay value at this time is set as tdc(n), where n ranges from 0 to 511.

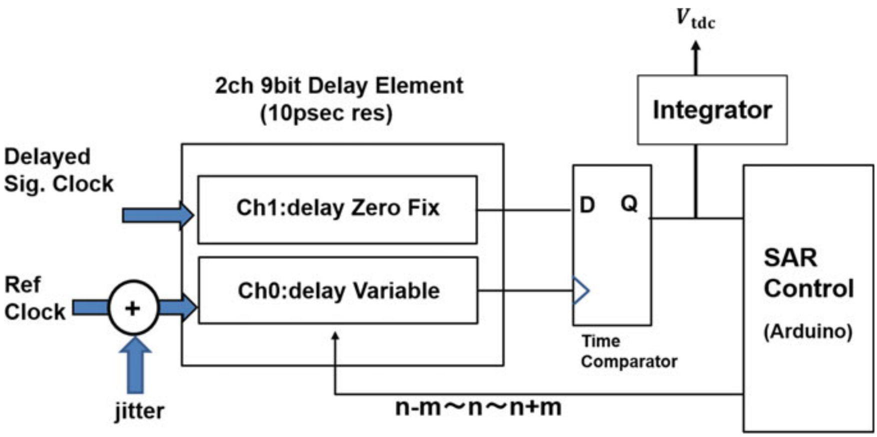


Fig. 3 Our prototype SAR TDC with jitter injection at the input

2. The voltage V_{tdc} is obtained by integrating the time comparator (DFF) output as shown in Fig. 3. Let the output voltage of tdc(0) be $V_{tdc}(\text{min})$ and the output voltage of tdc(511) be $V_{tdc}(\text{max})$, and normalize to $V_{tdc}(\text{min}) = 0$ and $V_{tdc}(\text{max}) = 1$.
3. Apply jitter generated from a signal source to ref clock, set tdc $= n$ measured in the above step 1 and measure $V_{tdc}(n)$. Vary n (from $n - m$ to $n + m$) by jitter application width and measure $V_{tdc}(n \pm m)$ at each code. Each V_{tdc} is also normalized from 0 to 1.0.
4. Jitter is applied to the ref clock, so that the delay value obtained by the SAR TDC is varied by the normally distributed probability density function in Fig. 4 (top).
5. The integral output V_{tdc} of the time comparator in each code corresponds to a cumulative distribution function between 0 and 1.0 as shown in Fig. 4 (bottom). Based on the cumulative distribution function, plot the $V_{tdc}(n - m)$ to $V_{tdc}(n + m)$ measurement data, and from the approximate curve, the point where the cumulative distribution $= 0.5$ is obtained as the most probable delay value [11].

Notice that the resolution of the variable delay is 10 ps, so the jitter application width is estimated to be around 100 ps. The accuracy can be improved by increasing the integration time of the time comparator.

3 Implementation and Measurement Results

This section shows implementation and measurement results of the proposed TDC using discrete electronic components on a board.

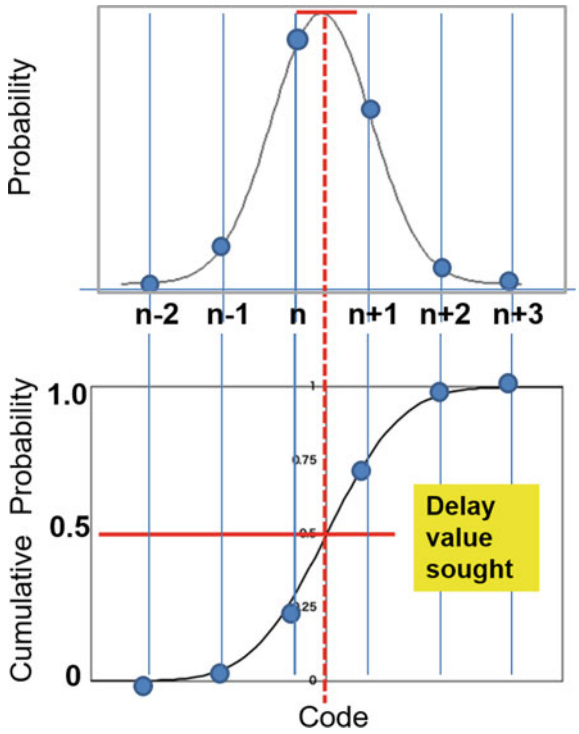


Fig. 4 Probability density function and cumulative distribution function

3.1 Variable Delay Device

In our experiment, a delay is generated between channels by controlling a dual 9-bit variable delay device (NB6L295M: ON Semiconductor) with Arduino device. We evaluated its linearity with a 2 GHz digital oscilloscope (R&S MSO-1022). The delay range is 11.13 ps/LSB between 0 and 5.6 ns with 9-bit resolution. After correcting the zero offset delay, the delay was calculated to obtain INL using the 0-to-510 end-point method.

Arduino was used in Raspberry Pi Pico development environment for the prototype TDC measurements.

The circuit configuration including the delay IC is shown in Fig. 5, while the measurement on the actual device is shown in Fig. 6. Here, the output of the function generator is set to SYNC. In addition, the delay IC input section level conversion (ADN4665) and output section level conversion (100EPT23) are chip-separated at CH1 and CH2 (linearity would be deteriorated if they were on the same chip).

We see from the measurement results in Fig. 7 that the nonlinearity of the device is estimated to be within around 1–2 LSB (~20 ps). The reason for the deterioration of linearity at high bits (around code 500) would be an error caused by the expansion of the measurement range of the digital oscilloscope.

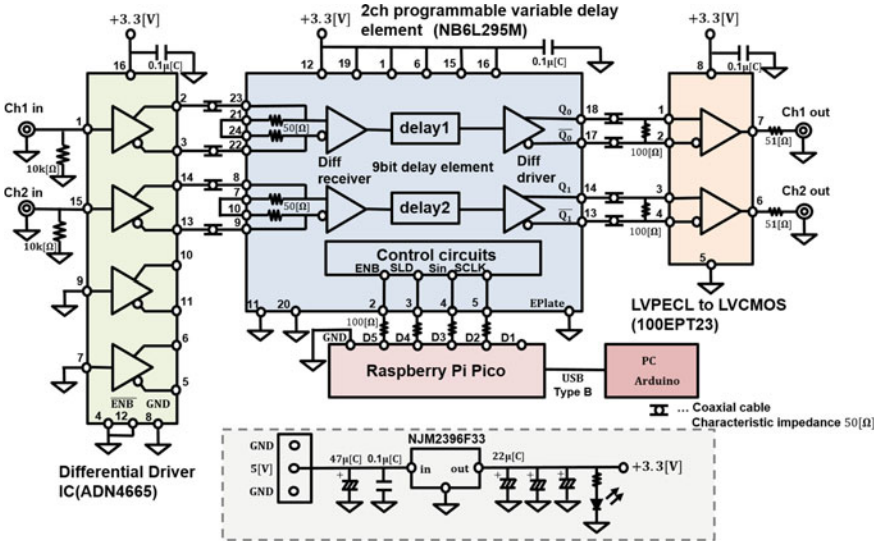


Fig. 5 Delay IC circuit using 9-bit programmable variable delay elements

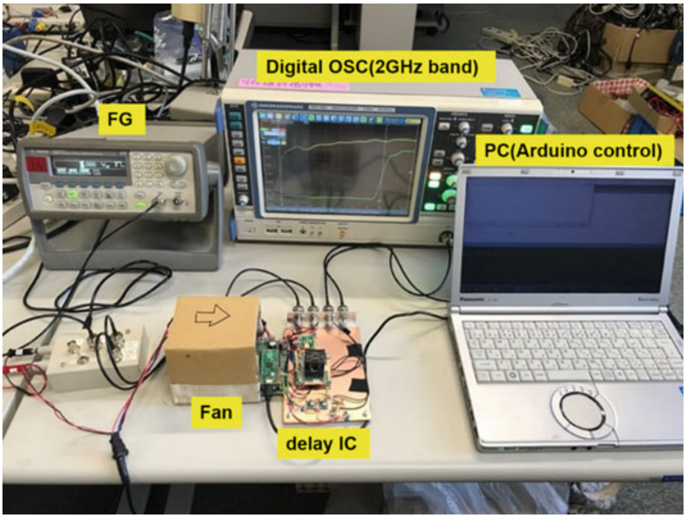


Fig. 6 Measurement environment for the delay IC circuit with 9-bit programmable variable delay element

3.2 SAR TDC

The linearity of the SAR TDC part using the delay IC measured in Sect. 3.1 is evaluated. The SAR TDC is shown in Fig. 8, and its measurement environment is shown in

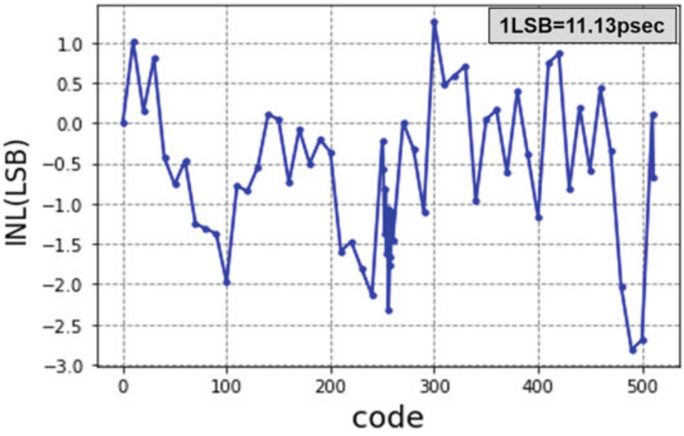


Fig. 7 Measured nonlinearity of the programmable delay line device

Fig. 9. Averaging was performed to reduce TDC measurement variation; averaging is necessary depending on the amount of jitter contained in the measurement signal. The measurement time was 260 μs without averaging.

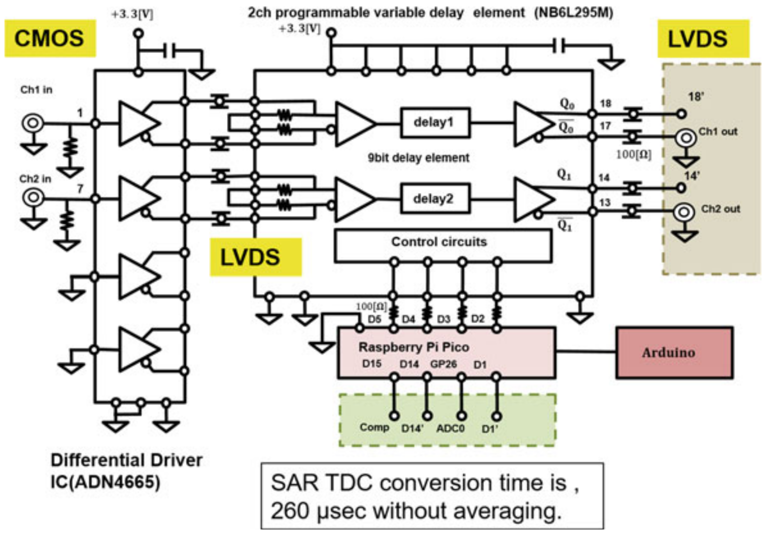
Multiple signal levels (CMOS, LVDS and PECL) are mixed and there are many level shift sections. The delay IC signals uses LVDS (differential signals), and the control signals are at CMOS level, while the input and output of the fixture are at 3.3 V CMOS level. This configuration was due to usage of available ICs for immediate delivery.

Measurements were carried out with averaging of 100 times. The measurement results in Fig. 10 show that the nonlinearity is within ±2 LSB.

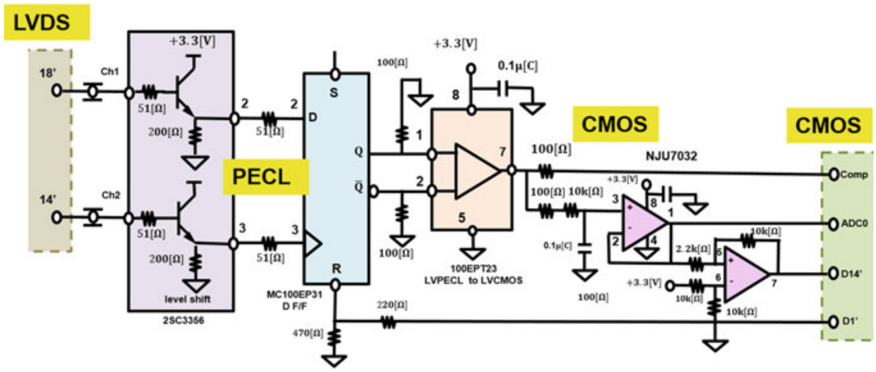
3.3 Vernier TDC

The fine delay values for out prototype TDC were evaluated using the Vernier principle presented in Sect. 2.1. In this experiment, the Levenberg–Marquardt method was used to obtain an approximate curve from the measured data [11]. This method is an iterative method that finds the minima of a function expressed in the form of the sum of the squares of a nonlinear function. This is considered as a combination of the steepest descent and Newton's methods and this is the standard method for solving a nonlinear least squares problem, which are to find the minimum of the sum of squares of a nonlinear function). From the approximate curve of the obtained cumulative distribution function (Fig. 11), the delay value of this cumulative distribution function is obtained as 0.5. As a result, a Vernier TDC value $N = 134.9$ (equivalent to 1 ps resolution) is obtained.

The Vernier TDC conversion time in the current experiment is about 13 ms, which consists of 2.6 ms for the upper 9 bits with 10-time averaging of the SAR TDC output



(a) Delay generation part



(b) Comparator and peripheral part

Fig. 8 SAR TDC part using the programmable delay line in Fig. 5 as a delayed input

to obtain the center code value and 10 ms for the lower 3 bits with 10 measurement points (jitter application of about 100 ps p-p) to obtain the approximate cumulative density function. It is estimated that measurement in a few milliseconds is possible by optimizing the measurement time, (e.g., by minimizing the time constant of the comparator).

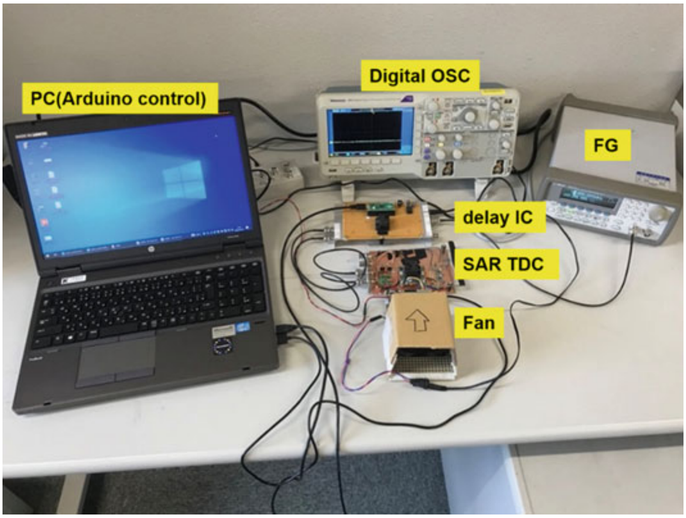


Fig. 9 Measurement environment of the prototype SAR TDC in Fig. 5

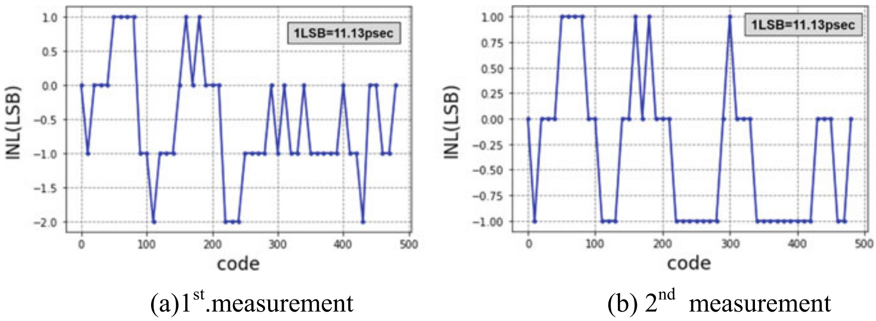


Fig. 10 Linearity assessment of SAR TDC

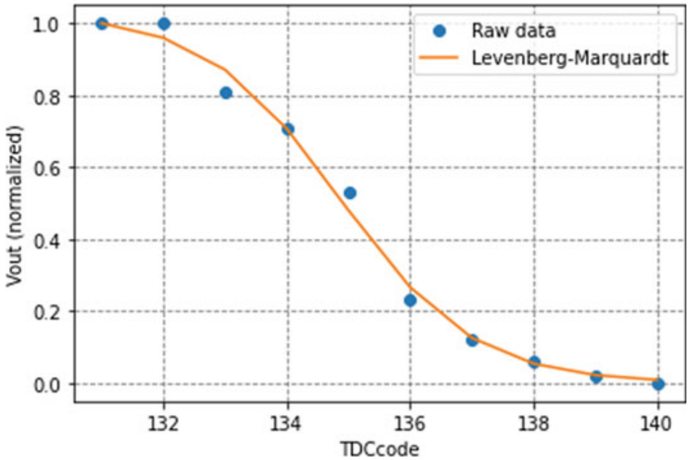


Fig. 11 Approximation curves for measured data using the Levenberg–Marquardt method

4 Conclusion

We have developed a 12bit 1 ps time resolution TDC using discrete circuits for LSI test system application; the high performance TDC can be implemented at low cost and in short development time. The prototype TDC was realized with combination of commercially available electronic components and standard analog modules. As a delay element used for obtaining the upper 9 bits with SAR TDC configuration, a 9-bit 10 ps resolution programmable variable delay element was controlled by Arduino and its INL was measured to be within ±2 LSB. For obtaining the lower 3 bits, the time resolution was improved by 10 times (up to 1 ps) by applying jitter, 100 times measurements and statistical processing.

The following are potential applications of this technology.

1. Accurate measurement of the pulse width time as well as the time difference.
2. Application of this technique to the ADC performance improvement, such as monotonicity and differential nonlinearity (DNL) improvement. Here the Vernier TDC time resolution is improved from the approximate cumulative density function and this is applicable to the ADC.

Future plans are as follows:

1. Application of an appropriate Gaussian characteristic jitter from the signal source to improve fine time measurement accuracy with variable jitter amount.
2. Usage of a coaxial phase variable (CDX-PS200-6GT, continuously variable from 0 to 200psec) to improve delay accuracy.

These are ready as shown in Fig. 12.

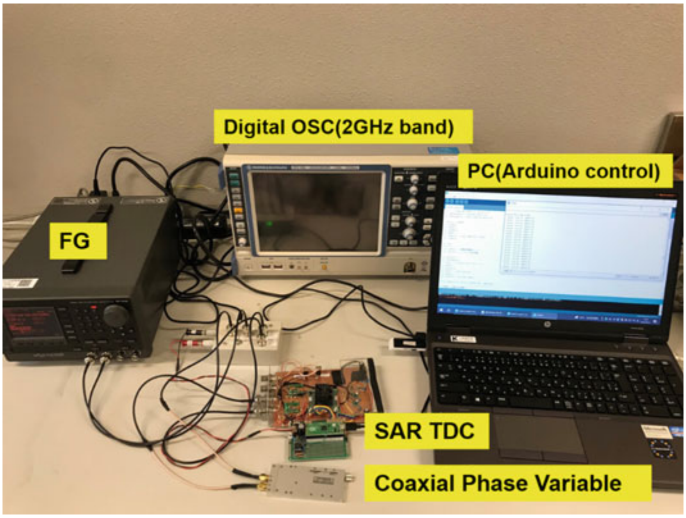


Fig. 12 Improved Vernier TDC circuit

References

1. Kobayashi H, Kuwana A, Wei J, Zhao Y, Katayama S, Tri TM, Hirai M, Nakatani T, Hatayama K, Sato K, Ishida T, Okamoto T, Ichikawa T (2020) Analog/mixed-signal circuit testing technologies in IoT era. In: IEEE 15th international conference on solid-state and integrated circuit technology, Kunming, China
2. Prasad KH, Chandratre VB, Saxena P, Pithawa CK (2011) FPGA based time-to-digital converter. In: Proceedings of the DAE symposium on nuclear physics, vol 56, pp 1044–1045
3. Arai Y, Baba T (1988) A CMOS time to digital converter VLSI for high-energy physics. In: IEEE symposium on VLSI circuits
4. Machida K, Ozawa Y, Abe Y, Kobayashi H (2018) Time-to-digital converter architectures using two oscillators with different frequencies. In: IEEE Asian test symposium, Oct 2018
5. Sasaki Y, Kobayashi H (2018) Integral-type time-to-digital converter. In: IEEE international conference on solid-state and integrated circuit technology, Nov 2018
6. Nelson M (2000) A new technique for low-jitter measurements using equivalent-time sampling oscilloscope. In: Automatic RF techniques group 56th measurement conference, Dec 2000
7. Ozawa Y, Ida T, Jiang R, Sakurai S, Takigami S, Tsukiji N, Shiota R, Kobayashi H (2017) SAR TDC architecture with self-calibration employing trigger circuit. In: IEEE Asian test symposium, Nov 2017
8. Lee J, Moon Y (2012) A design of vernier coarse-fine time-to digital converter using single time amplifier. J Semicond Technol Sci 12(4)
9. Jovanović GS, Stojčev MK (2009) Vernier's delay line time–to–digital converter. Scientific Publications of the State University of Novi Pazar. Ser A: Appl Math Inform Mech 1(1)
10. Jiang R, Li C, Yang M, Kobayashi H, Ozawa Y, Tsukiji N, Hirano M, Shiota R, Hatayama K (2016) Successive approximation time-to-digital converter with Vernier-level resolution. In: IEEE international mixed-signal testing workshop, July 2016
11. Ranganathan A (2014) The Levenberg-Marquarrdt algorithm. In: Tutorial on LM Algorithm 11.1, UC Santa Barbara, June 2014

Society 5.0 A Vision for a Privacy and AI-Infused Human-Centric Society Driving a New Era of Innovation and Value Creation

Elizabeth Koumpan and Anna W. Topol

Abstract In 2016, government of Japan defined Society 5.0 as a *human-centered societ*y (What is Society 5.0—Government of Japan: https://www8.cao.go.jp/cstp/english/society5_0/index.htmlCebit. Society 5.0: Japan's digitization. http://www.cebit.de/en/news-trends/news/society-5-0-japans-digitization-779 [1]) in which the need to address social problems and economic advancement is balanced. Based on innovation and a highly integrated physical and cyberspace, this optimized organizational structure enables the creation and delivery of new services and products for those who need them when they need them, providing unique value, breaking the sense of stagnation, and enhancing society as a whole. We already see the emergence of connected industries collaborating innovatively to deliver new services. Advances in edge technology, blockchain, and 5G provide means for more connected ecosystems. In the future, they will further focus on value creation for society, sharing the data, and utilizing AI, while keeping data protected and secure. This paper studies some of the challenges which need to be addressed to realize the Society 5.0 vision.

Keywords Society 5.0 · Ecosystems · Innovation · Connected industries · Human-centric · AI · Edge computing · Value creation · Data sharing

1 Introduction

Society 5.0 is a vision for privacy and an AI-infused human-centric society in which economic advancement is as significant as resolving societal problems. Its focus is to deliver a new added value to society by utilizing advanced technology like artificial intelligence (AI), edge computing, the Internet of Things (IoT), and 5G solutions

E. Koumpan (✉)
IBM Consulting, 3600 Steeles Ave East, Markham, ON L3R 9Z7, Canada
e-mail: ekoumpan@ca.ibm.com

A. W. Topol
IBM Research—Watson, 1101 Kitchawan Road, Yorktown Heights, New York 10598, USA
e-mail: atopol@us.ibm.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_19

[3–6]. Connecting people and organizations via access to data and technology will unlock the potential of data by sharing it safely and with privacy protection across connected ecosystems and industries.

The Covid-19 pandemic has accelerated Society 5.0 and ecosystem thinking, highlighting issues related to operational focus on efficiencies without appropriate risk management [7]. Addressing pandemics and other societal calamities requires collaboration beyond the public sector involving enterprises, subject matter experts (SMEs), and coordinating non-governmental organizations (NGOs) to drive new data gathering and management processes. How people work, learn, buy, and how businesses interact with their consumers, partners, and one another will be forever changed.

As we transition toward to Society 5.0, connected industries bring to the forefront new challenges and opportunities [8]. We already see how digital technologies are changing consumer expectations and blurring traditional industry boundaries. Uber Technologies changed Transport and Logistics Industry, while Apple, a Technology Company, now offers a credit card [9]. Ecosystem expansion of Ping An of China is bringing together health care, insurance, housing, and banking. Amazon's business transitioned from books to retail and beyond, creating the industry notion of "getting Amazoned" [9, 10]. We are also observing the creation of new ecosystems that become multi-sided value nets with enterprises at their core. These ***ecosystems focus on access to ext***rinsic organizations. They use data and AI to create new value via analytical insights [11].

2 Business Trends and Drivers for Society 5.0

The Society 5.0 trends will cause new business models to shift from output based (i.e., buy/sell/own for profit) to outcome/impact based. In this new model, focus will be on new personalized, purpose-led services involving ecosystem participants from multiple industries, driving higher incomes for participants and other businesses, the stickiness of participants in the ecosystem, decreasing the cost of acquiring customers, etc. The first question we address: ***What are prevailing business trends and drivers that are sufficiently disruptive and consequential for delivering on the promise of Society 5.0?***

Figure 1 Top 10 transition trends and new ecosystems drivers highlight our findings, as identified based on existing work done by IBM helping clients with human-centered ecosystems and analyzing examples in public domain. In summary, for Society 5.0, trust and human centricity will lead to advancements in:

- **Ethics, Impact, and Purpose**—Open, trusted, peer-endorsed services/ products.
- **Decentralization of Control/ Power**—more loosely coupled ecosystems where leaders release more power to participants to fuel the "network" effect.
- **Data Democratization**—bring your own data, data used for social and sustainable innovation.

FROM SOCIETY 4.0	Society 4.0 EXAMPLEs	TRANSITION TREND & Ecosystem Drivers	TO SOCIETY 5.0	Society 5.0 EXAMPLEs	Implic
Centralised/ Enterprise at Core Predictable Top Down Structure Few Leading Players Demand Anticipated. Scale visible early Fragmented pre-determined end to end CX (device, UI, QoS)	Apple	**DECENTRALISATION of control and power**	**Decentralised** Proxies, Intermediaries, New Players. VIRAL, Organic/ Fragmented Loosely Coupled; Bottom Up Structure; Many Diverse Participants. Seamless, assembled end to end CX (device, UI, QoS)	YARA	- Moving to a “N networks” archi - extensible, conf opt-in/opt-out s based on frictio exchanges. - Fragmented Da Data Flows / Da Data in Motion
Traditional Industry Verticals Co-Exist Industry Verticals Digitised & Reinvented in Verticals: Industry 3.0 –> 4.0	AirBnB, UBER,	**INDUSTRY REFORM**	**Converge around “macro-systems”**: MOBILITY system, ENERGY System, FOOD System, HOME/ City Systems etc Industry 4.0 –> 5.0. New standards to open more participation in closed industries.	HONDA 5 Strategy Mobility Systems (Auto, Energy, O&G, OEMs)	- Industry Conve displaces and disintermediate legacy industrie - Regulations and evolve to open participation an Data Exchanges - Industries Conn Converge aroun systems”: Mobi Food, Home/
Enterprise at Core - Digital Explosion Dispersed, Fragmented, Siloed, Proprietary Fragmented experiences. Disproportionate use of data for monolithic centralized ML	Visa	**CONNECTED SYSTEMS**	**Society at Core** - Connected Cyber/ Physical Society, Mesh/ Contextual computing Seamless experience. System of systems prevail - Highly Adaptive Cyber Physical Systems –rise of zero touch, frictionless Context aware algorithm zoo at the edge	BOSCH Connected World Tradelens	- Connected Cyb Society will be driven by the instrumentation physical world Edge Computin - Multiple types Interoperability of the stack. No New standards protocols. Mov “search to “recommendati - Data movemen mobility and da interconnection control and secu governance cha
Consumer Customer Deterministic ‘top down’ bilateral incentives	Alibaba Amazon	**PROSUMER SHIFT**	**Prosumerization & Enterprization of the consumer / SMB** Creator/Influencer: Bespoke services, Free-range, digital amniotic evolved microproducts: Gig Economy, P2P	BBOX, Shopify, Equigy Simple Bank (BBVA)	- A prosumer is a consumes and p product. - Empowered use ownership, so t monetise & sha driven economy - This will lead t proliferation of products and mi that in turn will influencer and g
Output based Product/ Goods/ Service Transactions Economics based on Buy/ Sell/ Own for Profit	BMW (sell cars) Sell Products	**OUTCOMES Value Creation / Value Capture**	**Outcome based** Adaptive ‘bottom up’ ecosystem first incentives Intangibles/ Products/ Goods/ Services Exchanged	BMW (mobility) Rolls Royce (Power by the hour)	- Trusted collabo transparency & in new systems - Data is needed trace, track, aut certify

Fig. 1 Top 10 transition trends and new ecosystems drivers (this study)

- **Connected Cyber/Physical Society** -the instrumentation of the physical world with IoT and edge computing.
- **New Data Sources and Standards** will combine existing datasets with new ones to set the foundation for contextual computing and highly adaptive cyberphysical systems for many industries.

- **Resiliency by Design**—a guiding design principle that is not only a technology requirement but also a business imperative that will create opportunities for new entities like "**Group Formed Networks**" (based on shared interests)

3 New Ways to Monetize and Create Value

To date, enterprise valuation and monetization models focused on "enterprise-centric" views are not sufficient for the evolving ecosystem value nets emerging in Society 5.0. They need to go beyond enterprise-centric value models and look at the industry and societal value nets that form the heart of sustainable ecosystem success. As Society 5.0 evolves, the decentralized ecosystems will become the foundational source for value creation and distribution. That is why we see a strong need to develop methods and approaches to understand growth and dynamics of complex ecosystems. A complex multi-sided monetization strategies will be adopted, further increasing the merge of industry boundaries with producers and consumers as stakeholders side-switching and playing either role at any given time [12].

An additional complication arises from the fact that supply chain economics today is not reckoning values with unaccountable profit related to the quality of life, strengthening community, the health of our environment (air, soils, water, plants, and minerals), or building trusting relationships. These natural capital assets provide clean water and air, a supply of food, medicine, shelter, energy, and raw materials for manufacturing. They also deliver more discreet benefits, including climate regulation, natural disaster control, pollination, and, therefore, the production of crops. They are critical for Society 5.0 as they related to securing the necessities of life for farms, communities, organizations, and ecosystems in general [13]. Figure 2. Enterprise vs. ecosystem value nets highlights the key transition aspects from the traditional enterprise value model (based on capital invested in the organization, information and technology, people, and processes) to the ecosystem value nets (based on natural, infrastructure, social, cultural, collective, and individual capital components).

To fully monetize in the ecosystem economy, understanding the value flows (which tightly link to data flows) helps to determine points of "settlement," "trade," and "value" at microlevels. This is essential when identifying the complex structure of multi-stakeholder remuneration. Figure 3. Enterprise versus ecosystem value nets, showing the differences between the enterprise value chain and exemplary ecosystem value net assessed in our study. To do so, full process of the data flow across the ecosystem will have to be understood, including its generation, usage, storage, ownership, and elimination. To establish new industry structures and further advance AI deployments, enablement of cross-industry information sharing platform supported by a common and shareable by all infrastructure is critical [14].

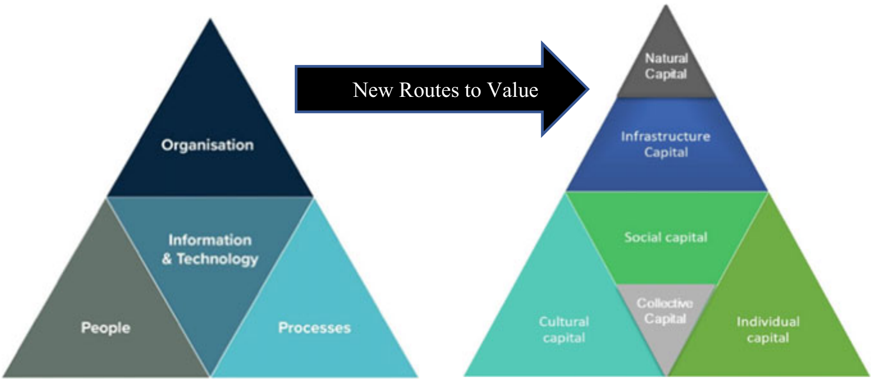


Fig. 2 Enterprise versus ecosystem value nets

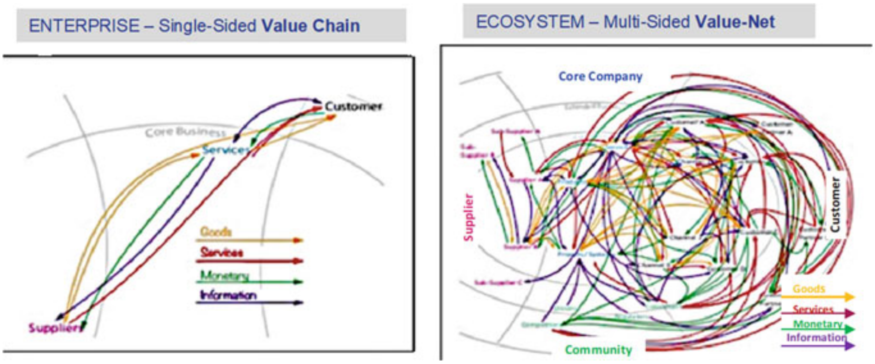


Fig. 3 Enterprise versus ecosystem value nets (this study)

4 Utilization of Data and AI

We live in times when data has become a commodity [15]. *What is done with the data* leads to value creation, enabled when an organization institutes data democratization, coupled with appropriate data governance (see Fig. 4. Utilizing data for new insights).

The nature of the data involved (source and quality) and the process of data use (how and where data us used) define the data value opportunity. As shown in Fig. 5. Routes to ecosystem value capture, in 2017, IBM identified six routes to capture value in the ecosystem: Anchor Enterprise, Anchor Ecosystem, Standards Route, Peer to Peer, Data Led, and Capability/Asset/Service Led.

The real value of data within an ecosystem and the key to unleashing the next level of value creation in Society 5.0 span from a network that leverages data input from multiple entities and sources. ***User-centric data ecosystems are becoming a norm, where participants can collect and share data, collaborate, exchange information, and deliver insights.*** Emphasis on ease of consumption and understandable access

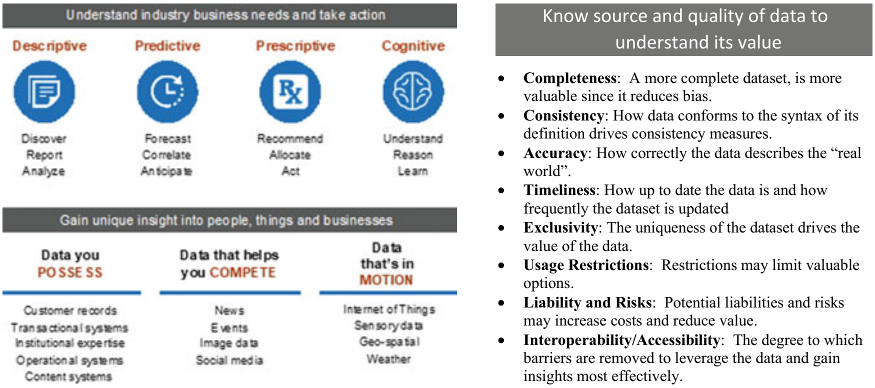


Fig. 4 Utilizing data for new insights

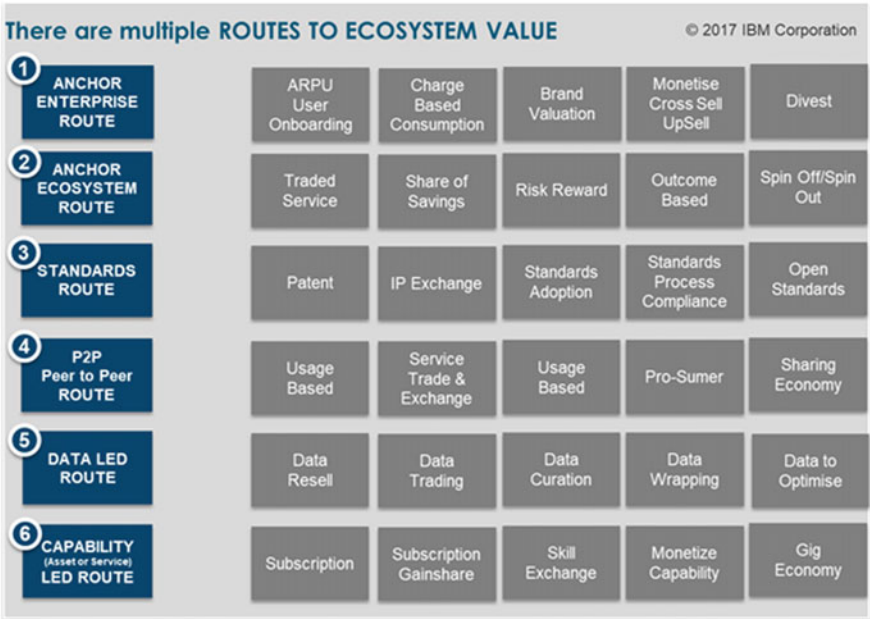


Fig. 5 Routes to ecosystem value capture (IBM Corporation 2017)

to good data by all participants expedites end-user adoption of and engagement with the ecosystem and hastens the value creation process.

Of course, not all ecosystem participants will want to openly share all data. AI models are valuable intellectual property and may or may not directly have access to the data they were trained on. Further, regulations may restrict data sharing and require the enforcement of privacy and licensing rights.

Trusted data and metadata are vital when developing and running AI-based solutions. It requires appropriately curated sources (such as non-biased data) and appropriate data governance (standards, provenance, and proper use).

4.1 Data Value Realization

Sometimes called "decision-centric computing," [16] the need to understand and utilize data goes beyond data integration and governance. The process of gaining insights and realizing full potential from data consists of three steps: (1) putting data into context, (2) understanding its relationships with other data or events, and (3) taking action or decision. In Society 5.0, this will be enabled by advanced data sharing patterns among participants of the connected ecosystems with AI as a mechanism to extract the correct value from the shared data.

The Society 5.0 value proposition [17] rests on a logical framework in which several technology components interact to provide value to society, whether or not those members currently derive value from their participation in technology-oriented ecosystems. Such a logical architecture supports the function of new, more inclusive, higher value ecosystems by defining the content and relationships among a set of constituent physical architectures and, by extension, the parties affected by those ecosystems. In this human-centered society, value is realized via information sharing, infusion of AI capabilities, and elimination of geographic and language gaps to enable economic growth and personalized products and services [18]. Some of the early industry examples focused on collaborative systems include:

- In Oil and Gas—7 companies of Petroleum Association of Japan and 13 companies [19] of Petrochemical Industry Association jointly developed prediction models for internal and external corrosion—sharing plant data and verification of effects through analysis model collaboration
- Ship data center—to create an environment in which the use of navigational data is promoted by establishing and managing data related to ships, including the collection & accumulation of navigation data.
- Petroleum (O&G) data center—adapted to refinery safety by sharing various data and physical models on the platform.
- Universal material incubator—by classifying and analyzing the diverse technologies and peripheral information of dispersed materials and chemical industry by business and project, new materials can be created.
- Agriculture—as shown in Fig. 6. Crop farming cloud ecosystem example.

As assets and/or their corresponding data are being shared and exchanged securely among ecosystem participants or algorithms must be developed to follow data exchange standards and value the data in context based on the use case.

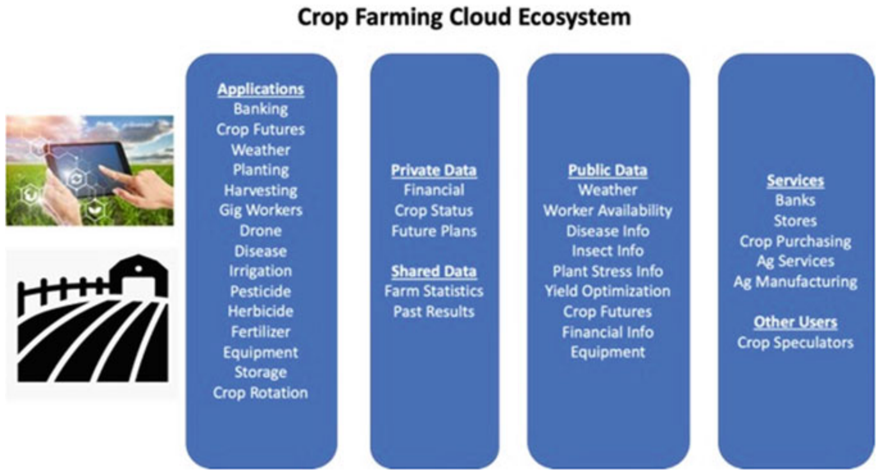


Fig. 6 Crop farming cloud ecosystem

There are many advancements related to ability to share information without sharing the data. These include:

- Federated learning: in which each participating organization trains an ML model on using their own data. Then the models are combined as federated ML model.
- This aggregation enables the participants to develop a better model without sharing or access to other participants' data.
- Differential privacy: focuses on withholding information about individuals by only sharing information about dataset groups.
- Fully Homomorphic Encryption (FHE): focused encryption that permits operations (computations) directly on encrypted data, without first decrypting it.
- Zero Knowledge Proofs (ZKP): Method by which Prover can prove to a Verifier that a given statement is true (or a given value is known) avoiding exchange of any additional information.
- Secure Multiparty Computation (SMC): Cryptographic protocol sharing a computation (an aggregate function) without sharing data or details (e.g., the ecosystem calculates number of parts in its inventory without sharing any individual company's inventory).

4.2 Value Distribution

The value of the data is limited by the expected increase in profits for the organization to the extent that it changes the business (economic) value of the desired outcome. That change can be accomplished by decreasing the likelihood of "bad" outcomes (reducing risk) or increasing the probability of good outcomes and needs to take the ecosystem's participant's risk profile into account. Decision theory and modern utility theory [20] offer a mathematical method for modeling risk tolerance. Generally, as wealth increases, so does risk tolerance. Also, economic and market forces dictate how value is distributed among the participants. Our social media posts generate income for the social media companies, but only "entertainment" for us.

Key question for the Society 5.0 ecosystems of partners is: How much of the excess value created using the data goes to each participant or the ecosystem as a whole? Do participants who benefit more pay more? Does everyone pay the same, and how they derive value is up to them? Do we simplify the problem by creating different classes of participants? For example, different costs for small family farms versus large corporate farms for sensor data? Who has the market power? For instance, in an ecosystem run by a corporation, like the Apple Store or Amazon Marketplace, the ecosystem owner has enormous market power and can almost dictate prices and retain a significant proportion of the transaction value as an enabling fee.

The three typical approaches typically used to value any asset can also be applied to data valuation [21]:

The Income Approach is the most theoretically staunch however requires a subject matter expert (SME) to deploy. It takes into account the ability to generate future cash flows for the data owner and the incremental cash flow benefits to the consumer who is using (or purchasing) the data.

Market Approach takes into account the current market value for this type of data, considering the most relevant business and technology substantiation.

The Cost Approach typically focuses on current market value (not future state) and, when considered from the data provider's viewpoint, accounts for the total cost to generate the data assets (including design, test, and delivery). However, from a data consumer point of view, the cost approach accounts for (1) the cost to reproduce or (2) the cost to replace the similar data input and its generation method.

A completely different approach to valuing intangible data assets is based on **the Cooperative Game Theory Approach** [22]. In this approach, there are N "players" who can form a coalition S (a set of players). All the members of a coalition cooperate with each other to increase the value of the coalition. The *grand coalition* is the coalition of all players. A valuation function *v(S)* computes the "value" of the coalition *S*. A function *v(S)* is defined for all subsets *S* of the grand coalition. A solution is a vector *x(i)* that represents the allocation of rewards or values to each member "i" of the coalition. For our purpose, each player is a dataset or capability, and a coalition *S* is the aggregation of datasets and capabilities, creating a single, combined dataset. The *x(i)* is the "value of a dataset *i*." We train a specific ML algorithm *M* using all the data in *S* (against fixed ground truth if supervised ML) to create model *M(S)*. It is a very general approach. Valuation function *v(S)* can be defined in many different ways.

5 Next Steps

There are technical, business, regulation, and corresponding policy challenges related to data, information, and infrastructure sharing, especially in the view of future complex multi-sided ecosystems. Massive efforts of AI engineering (similar to traditional software engineering) focus on ensuring the quality of data, metadata, and models, guaranteeing fairness and addressing explainability. In addition, legal relationships between AI-based software right and liability are being considered. Data shared by users may have economic value to users and vendors but clearly defined licensing, regulations, and understanding of how such data is used are required. Deployment of advanced technology for the protection of such data is also critical.

As user's data, information or assets are exchanged in multi-ste ecosystems; the question arises on how to track and measure the value it contributes to all the ecosystem participants. ***Value frameworks* need further refinement to be more sensitive and reflect the natural, social, and cultural capital from human interactions with ecosystems.** They need to go beyond accounting for economic impact and include unprofitable externalities (e.g., enhanced quality of life, community empowerment, or care of our environment). These traceable value nets will automatically quantify the value gained by the contributor from sharing its assets to ensure a fair exchange of value.

Acknowledgements We are very thankful to Xinlin Wang, Richard Hopkins, Marshall Lamb, Farzaneh Ghods, Adrian Papaccia, Cynthia S Unwin, Matt Seul, Nitin Gaur, Ted Tritchew, Scott Gerald, and Bill Chamberlin for their feedback and comments to the draft of the AoT study. We are grateful to the IBM Academy of Technology, Seth Dobrin, and Teresa Hamid, who were the champions of this study.

References

1. What is Society 5.0—Government of Japan. https://www8.cao.go.jp/cstp/english/society5_0/index.htmlCebit. Society 5.0: Japan's digitization. http://www.cebit.de/en/news-trends/news/society-5-0-japans-digitization-779
2. Cabinet Office. https://www8.cao.go.jp/cstp/english/society5_0/index.html
3. Cabinet Office (Council for Science, Technology, and Innovation) Government of Japan, Comprehensive strategy on science, technology, and Innovation (STI) for 2017. https://www8.cao.go.jp/cstp/english/doc/2017stistrategy_main.pdf
4. Unesco (2017) Japan is pushing ahead with Society 5.0 to overcome chronic social challenges. https://en.unesco.org/news/japan-pushing-ahead-society-50-overcome-chronic-social challenges
5. Nahavandi S (2019) Industry 5.0–a human-centric solution. Sustainability 11(16):4371
6. Nirmala J (2016) Super Smart Society: Society 5.0. https://www.roboticstomorrow.com/article/2016/09/super-smart-society-society-50/8739
7. McKinsey & Company—Covid 19 Implication for Businesses. https://www.mckinsey.com/business-functions/risk/our-insights/covid-19-implications-for-business
8. Cairo Review—Society 5 and the future Economics. https://www.thecairoreview.com/essays/society-5-0-and-the-future-economies/

9. https://www.apple.com/apple-card/
10. The rise of the aggregators. https://www.occstrategy.com/media/1331/to-platform-or-not-to-platform.pdf
11. Japan Ministry of Economy Trade and Industry. https://www.meti.go.jp/press/2019/04/20190404001/20190404001-1.pdf
12. https://naturalcapitalcoalition.org/coalition-organizations/
13. https://frankdiana.net/2022/08/08/why-ecosystems-why-now/
14. Hitachi—Inspire the next. https://www.hitachi.com/rev/archive/2017/r2017_06/trends/index.html
15. Utilization of Data & AI. https://www.meti.go.jp/press/2019/04/20190404001/20190404001-2.pdf
16. Global Technology outlook. https://www.idc.com/getdoc.jsp?containerId=US43282917
17. https://www.japan.go.jp/abenomics/_userdata/abenomics/pdf/society_5.0.pdf
18. Koumpan E, Topol AW (2021) Promoting economic development and solving societal issues within connected industries ecosystems in society 5.0. In: Advances in artificial intelligence, software and systems engineering, pp174–183
19. Transforming Petroleum industry. https://www.paj.gr.jp/english/data/paj2015.pdf
20. Howard R (1997) In: Decision analysis: introductory lectures on choices under uncertainty. McGraw Hill. ISBN 978-0-07-052579-5
21. Personal Data Protection Commission Singapore (2019) Guide to Data Valuation for Data Sharing
22. Bagwell K, Wolinsky A (2002) Game theory and industrial organization. ch. 49, In: Handbook of game theory with economic applications, vol 3. pp 1851–1895

Systematic Review and Propose an Investment Type Recommender System Using Investor's Demographic Using ANFIS

Asefeh Asemi, Adeleh Asemi, and Andrea Ko

Abstract The development of investment recommender systems (IRSs) has increased due to advancements in technology. This study aims to present a new model for IRSs based on potential investor's demographic data and feedback, using fuzzy neural inference solutions. Both qualitative and quantitative methods were used in this research, including a review of past studies and analysis of data through Scopus analyze tool, Voyant, and VosViewer. The proposed model combines expert's knowledge with demographic data to present the most suitable type of investment through an adaptive neuro-fuzzy inference recommender system. The model is processed in several steps, including clustering investment data types in JMP and proposing the results through MATLAB. This study provides a framework for IRSs that can give relevant and accurate recommendations for potential and actual investors, thus enhancing their investment experience.

Keywords Adaptive neuro-fuzzy inference system · ANFIS · Investment recommender system (IRS) · Demographic data · Investment type · Investment product · Potential investors · Investor feedback · Investment service

A. Asemi
Doctoral School of Economics, Business, and Informatics, Corvinus University of Budapest, Budapest, Hungary
e-mail: asemi.asefeh@uni-corvinus.hu

A. Asemi
Department of Software Engineering, Faculty of Computer Science and Information Technology, Universiti Malaya, Kuala Lumpur, Malaysia
e-mail: adeleh@um.edu.my

A. Ko (✉)
Corvinus University of Budapest, Budapest, Hungary
e-mail: andrea.ko@uni-corvinus.hu

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_20

1 Introduction

Customer demographic data and information are one of the most important sources in examining the past, present, and future of a company. In recommender systems, these data and information play a very important role in providing suitable recommendations to the customer. Kanaujia and his colleagues also believe that recommender systems are tools based on customer needs [1, 2]. This issue is obvious and considering that the customer's demographic information plays a significant role in providing appropriate advice to the customer. This research has tried to provide a new model for the recommender system to recommend the type of investment by using adaptive neural fuzzy inference (ANFIS). For this purpose, the demographic information collected from potential customers is classified using machine learning techniques and considered as system input, and the system output is the suitable type of investment for the specific investor.

2 Literature Review

Several studies have explored the fundamental concepts of this research topic. The following are some examples of these studies. Paranjape-Voditel and Umesh [1] proposed a recommender system for stock market portfolios based on association rules that analyzes inventory records and suggests suitable portfolios. In 2017, a collaborative filtering-based recommendation device was proposed for financial analysis based on savings, costs, and investments using Apache Hadoop and Apache Mahout [2]. Hernández et al. [3] evaluated the state of the art in financial technology to design a recommender system. They introduced a social computing platform based on virtual organizations that allows people to enhance their experience in activities related to financing recommendations. Tejeda-Lorente et al. [4] developed a risk-aware recommender system for hedge funds, considering multiple factors such as current yields, historical performance, and industry-wide diversity. Faridniya and Faridniya [5] presented a resource allocation and investment selection model using data envelopment analysis for the Social Security Organization (SSO) in Iran, which is responsible for the state pension fund. Their study revealed that the SSO's current investment strategy for the pension fund is at risk of bankruptcy, and a continuation of this trend would increase the likelihood of bankruptcy. Tarnowska and colleagues [6] designed a recommender system to enhance customer loyalty, which supports managers in identifying effective strategies for decision-making. The system enables customers to view anonymous feedback from other customers and includes a sensitivity evaluation feature. Sulistiyo and Mahpudin [7] investigated the demographic factors that influence the choice of investment types among amateur golfers in Karawang City, Indonesia, dividing investment types into two categories: real estate and financial assets. Their research showed that demographic factors such as gender, occupation, education, number of family members, and income significantly affect the choice of investment type, while age does not have a significant impact. On March 4th, 2023, a search was conducted in Scopus, utilizing the following formula, to retrieve relevant documents on the research topic of interest:

> (TITLE ("recommender" OR "recommendation" OR "decision") AND TITLE (investment AND system)) AND (LIMIT-TO (PUBYEAR, 2023) OR LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015) OR LIMIT-TO (PUBYEAR, 2014)).

After the search, 154 documents were found, out of which 44 full records were imported into Zotero. Figure 1 displays the distribution of the documents by subject area, indicating that the majority of them belong to the field of computer science.

According to Fig. 2, the documents in IRS are categorized by their funding sponsors. The majority of the documents are supported by the National Natural Science Funds of China, followed by the Fundamental Research Funds for the Central Universities. The research projects supported by the Fundamental Research Funds for the Central Universities mainly focus on books and symposium papers.

After conducting an analysis, it was determined that there are 3947 data repository files in Mendeley Data related to "recommender" or "recommendation" or "decision" and investment and system, which are available in various formats such as dataset (1267), tabular data (639), document (366), collection (285), text (212), software/code (109), image (57), file set (41), slides (26), video (26), audio (2), geospatial data

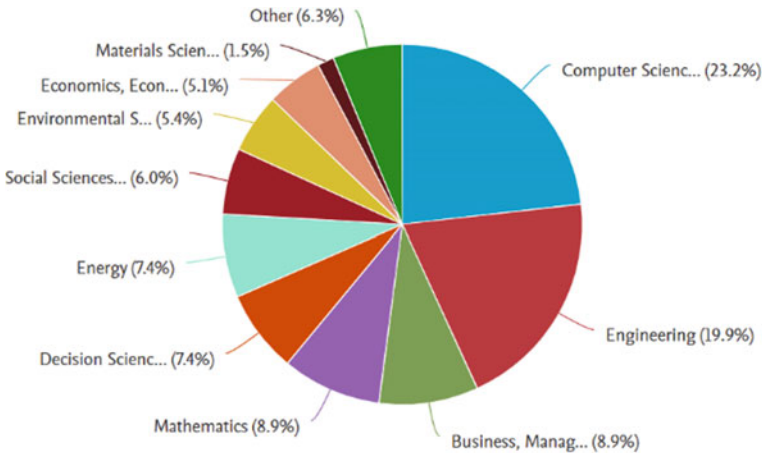


Fig. 1 Documents by subject area in IRS by scopus analysis tool

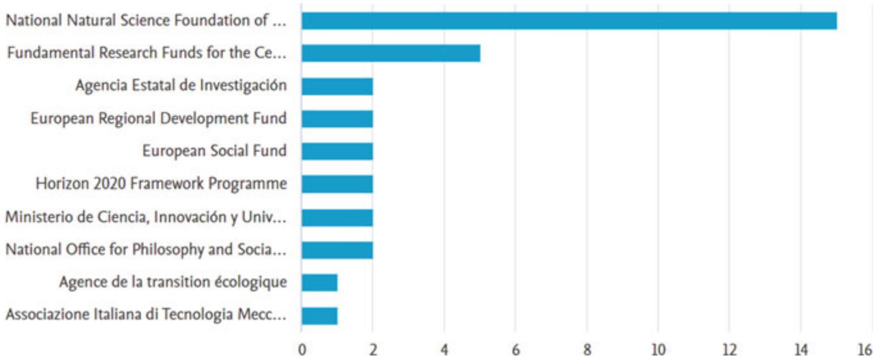


Fig. 2 Documents by funding sponsor in IRS by scopus analysis tool

(1), and others (2342). The author and index keywords from Scopus documents were analyzed using Voyant, revealing a total of 2640 words and 1509 unique keyword forms. The vocabulary density was calculated as 0.572, and the readability index was 68.616. Among the most frequently occurring words in the corpus were investments (94), decision-making (60), investment decisions (39), decision support systems (36), and multi (26). To provide a comprehensive overview of all keywords, a Cirrus was created and included in Fig. 3, which displays a keyword cloud view of the most commonly occurring Author and Index keywords. Additionally, Fig. 4 shows the co-occurrence of the retrieved keywords in Scopus (154 documents) using VosViewer.

VosViewer was used to investigate co-authorship patterns in a set of 154 documents retrieved from the IRS database. Figure 5 displays the co-authorship relationships among the documents, with most of them being attributed to Wang Y. The authors of these articles represent a diverse range of countries, including the United States, China, Turkey, and Poland, among others.

Fig. 3 Keyword cloud view of the most frequently occurring author and index keywords by Voyant

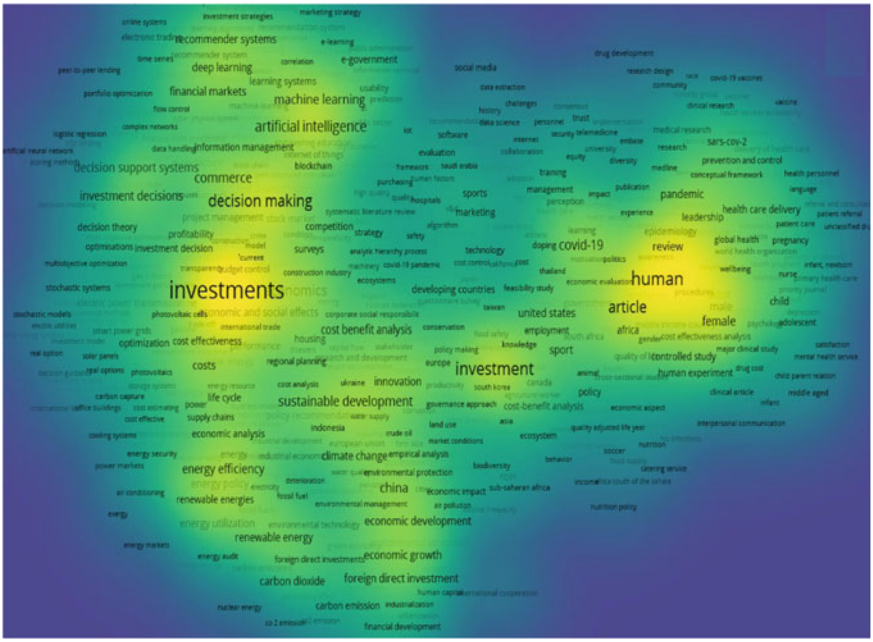


Fig. 4 Co-occurrence of the keywords in IRS by VosViewer

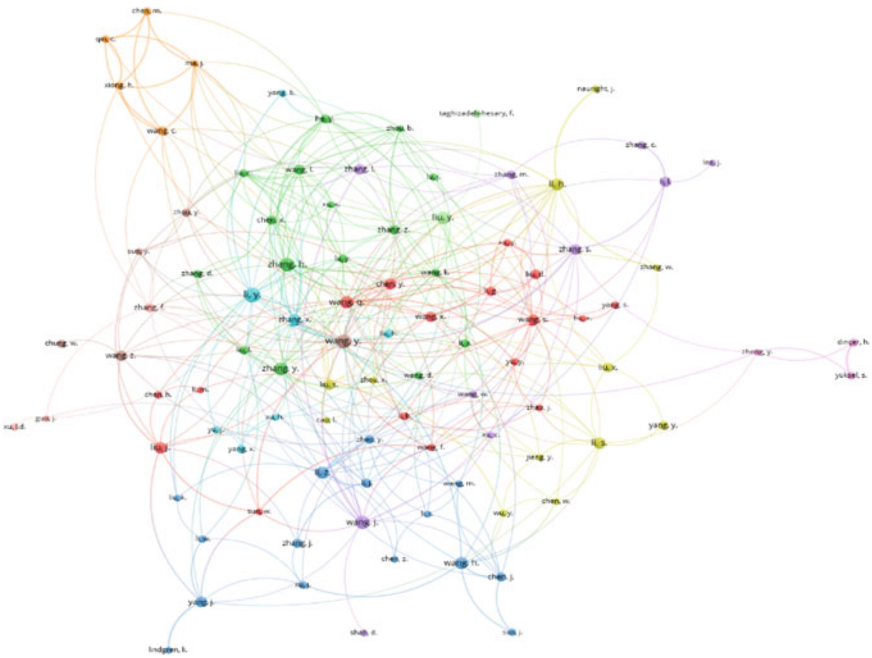


Fig. 5 Co-authorship in IRS by VosViewer

Table 1 Most frequent subjects in IRS (2019–2023)

Subject	Frequency
Decision support system	28
Renewable energy	9
Power systems	8
Artificial intelligence	6
Real estate	4
Multi-criteria decision analysis	3
Peer-to-peer lending	3
Machine learning	2
Carbon neutrality	2
System dynamics	2

Upon reviewing the previous research, it is evident that several topics gained popularity during the period from 2019 to 2023. These topics included artificial intelligence (AI) and machine learning, blockchain technology, renewable energy systems and investments, decision support systems and models, digital transformation and Industry 4.0, sustainability and green investments, big data analytics and predictive modeling, cybersecurity and risk management, peer-to-peer lending, and alternative financing models, as well as real estate investments and portfolio optimization. Table 1 displays the most frequent subjects based on the years 2019 to 2023 as per the IRS:

The most discussed topics include real estate investment, decision support systems, investment decision-making, and renewable energy systems. In 2020, decision support systems and real estate investment were frequently discussed, while in 2021, the focus shifted to renewable energy systems and investment decision-making. In 2022, renewable energy systems and decision support systems were still popular subjects. Looking ahead to 2023, the fintech ecosystem, decision support systems, and investment decision-making are expected to be the most talked about topics.

Table 2 displays the documents published in the last decade that have received at least five citations from the IRS. The document with the highest number of citations is Gottschlich and Hinz's 2014 paper. The data in the table has been retrieved from Scopus. While several studies have been conducted on this research, none have yet proposed a recommender system for investment products or types based on the demographic characteristics of potential investors. This study presents a novel ANFIS-based IRS that examines the demographic information of potential investors and, using expert opinions, recommends the most appropriate investment product for each customer group. This study's innovation is the collection of data from potential and existing investors to improve the ANFIS system. Additionally, the study utilizes an intelligent fuzzy framework to generate rules, which can be improved based on expert feedback.

Table 2 Most cited documents in IRS (2014–2022)

Authors	Title	Year	Citation
Gottschlich and Hinz [8]	A decision support system for stock investment recommendations using collective wisdom	2014	68
Salge et al. [9]	Investing in information systems: on the behavioral and institutional search mechanisms underpinning hospital's is investment decisions	2015	59
Zhou et al. [10]	Effects of a generalized dual-credit system on green technology investments and pricing decisions in a supply chain	2019	40
Starita and Scaparra [11]	Optimizing dynamic investment decisions for railway systems protection	2016	37
Ullah and Sepasgozar [12]	Key factors influencing purchase or rent decisions in smart real estate investments: a system dynamics approach using online forum thread data	2020	35
Kovačić et al. [13]	Optimal decisions on investments in urban energy Cogeneration plants—extended MRP and fuzzy approach to the stochastic systems	2017	33
Del Giudice et al. [14]	Real estate investment choices and decision support systems	2019	32
Geressu and Harou [15]	Screening reservoir systems by considering the efficient trade-offs—informing infrastructure investment decisions on the Blue Nile	2015	31
Yan et al. [16]	Pre-disaster investment decisions for strengthening the Chinese railway system under earthquakes	2017	28
Naranjo and Santos [17]	A fuzzy decision system for money investment in stock markets based on fuzzy candlesticks pattern recognition	2019	27
Fang et al. [18]	Assessment of safety management system on energy investment risk using house of quality based on hybrid stochastic interval-valued intuitionistic fuzzy decision-making approach	2021	24
Babaei and Bamdad [19]	A multi-objective instance-based decision support system for investment recommendation in peer-to-peer lending	2020	24
Lakhno et al. [20]	Development of the decision-making support system to control a procedure of financial investment	2017	24

(continued)

Table 2 (continued)

Authors	Title	Year	Citation
Teotónio et al. [21]	Decision support system for green roofs investments in residential buildings	2020	23
Mo et al. [22]	Delaying the introduction of emissions trading systems-implications for power plant investment and operation from a multi-stage decision model	2015	23
Kamari et al. [23]	A hybrid decision support system for generation of holistic renovation scenarios-cases of energy consumption, investment cost, and thermal indoor comfort	2018	22
von Appen and Braun [24]	Interdependencies between self-sufficiency preferences, techno-economic drivers for investment decisions, and grid integration of residential PV storage systems	2018	17
Renna [25]	A decision investment model to design manufacturing systems based on a genetic algorithm and Monte-Carlo simulation	2017	17
Flora and Vargiolu [26]	Price dynamics in the European Union emissions trading system and evaluation of its ability to boost emission-related investment decisions	2020	15
Ali et al. [27]	Does sustainability reporting via accounting information system influence investment decisions in Iraq?	2019	14
Kafuku et al. [28]	Investment decision issues from remanufacturing system perspective: literature review and further research	2015	14
Keding and Meissner [29]	Managerial overreliance on AI-augmented decision-making processes: how the use of AI-based advisory systems shapes choice behavior in R&D investment decisions	2021	13
Jankova et al. [30]	Investment decision support based on interval type-2 fuzzy expert system	2021	12
Ribas et al. [31]	A decision support system for prioritizing investments in an energy efficiency program in favelas in the city of Rio de Janeiro	2015	12
Akhmetov et al. [32]	Mobile platform for decision support system during mutual continuous investment in technology for smart city	2019	11

(continued)

Table 2 (continued)

Authors	Title	Year	Citation
Quitoras et al. [33]	Toward robust investment decisions and policies in integrated energy systems planning: evaluating trade-offs and risk hedging strategies for remote communities	2021	10
Akhmetov et al. [34]	Decision support system about investments in smart city in conditions of incomplete information	2019	10
Bruaset et al. [35]	Performance-based modeling of long-term deterioration to support rehabilitation and investment decisions in drinking water distribution systems	2018	10
Li et al. [36]	Risk decision-making based on Mahalanobis-Taguchi system and gray cumulative prospect theory for enterprise information investment	2016	10
Akhmetov et al. [37]	Model for a computer decision support system on mutual investment in the cybersecurity of educational institutions	2018	9
Cano et al. [38]	A strategic decision support system framework for energy-efficient technology investments	2017	9
Al-Augby et al. [39]	Proposed investment decision support system for stock exchange using text mining method	2016	9
Cabrera-Paniagua et al. [40]	A novel artificial autonomous system for supporting investment decisions using a Big Five model approach	2021	8
Tao et al. [41]	Review and analysis of investment decision-making algorithms in long-term agent-based electric power system simulation models	2021	8
Khalatur et al. [42]	Multiple system of innovation-investment decisions adoption with synergetic approach usage	2020	8
Papapostolou et al. [43]	Optimization of water supply systems in the water—energy nexus: model development and implementation to support decision-making in investment planning	2018	8
Siejka [44]	The role of spatial information systems in decision-making processes regarding investment site selection	2017	8

(continued)

Table 2 (continued)

Authors	Title	Year	Citation
Hu and Zhou [45]	A decision support system for joint emission reduction investment and pricing decisions with carbon emission trade	2014	8
Rühr et al. [46]	A classification of decision automation and delegation in digital investment management systems	2019	7
Ortner et al. [47]	Incentive systems for risky investment decisions under unknown preferences	2017	7
Li et al. [48]	Shared energy storage system for prosumers in a community: investment decision, economic operation, and benefits allocation under a cost-effective way	2022	6
Sun et al. [49]	Decision-making of port enterprise safety investment based on system dynamics	2020	6
Xue et al. [50]	Multi-scenarios based operation mode and investment decision of source-storage-load system in business park	2019	6
Thomas et al. [51]	A decision support tool for investment analysis of automated oestrus detection technologies in a seasonal dairy production system	2019	6
Mutanov et al. [52]	Investments decision-making on the basis of system dynamics	2018	6
Luo [53]	Application of improved clustering algorithm in investment recommendation in embedded system	2020	5
Wei et al. [54]	Joint optimal decision of the shared distribution system through revenue-sharing and cooperative investment contracts	2019	5
Ren and Malik [55]	Investment recommendation system for low-liquidity online peer-to-peer lending (P2PL) marketplaces	2019	5
Kozlova et al. [56]	New investment decision-making tool that combines a fuzzy inference system with real option analysis	2018	5
Niu et al. [57]	Improved TOPSIS method for power distribution network investment decision-making based on benefit evaluation indicator system	2017	5
Scaparra et al. [58]	Optimizing investment decisions for railway systems protection	2015	5

3 Methods

In this study, both quantitative and qualitative methods were used to develop a new model for the IRS. Specifically, machine learning and fuzzy inference logic techniques were implemented. To cluster investment types/products, K-means clustering, an unsupervised machine learning algorithm, was used with the JMP clustering toolbox. In this type of clustering, the number of clusters is represented by k, and the cluster of a data point is determined based on its distance to each cluster. The resulting clusters were used as inputs for an ANFIS fuzzy system, which was used to predict investment types based on customer demographic groups. The ANFIS system simplifies the establishment of logical judgment and follows Sugeno's rules, based on "IF_THEN" logic. The input membership functions were used to generate the fuzzy rules, and two categories of rules were considered: system-generated and manually created rules based on expert knowledge. Data for the study were based on responses to the Portfolio Investment questionnaire, with demographic characteristics and investment types experienced by the respondents used as variables. The dataset included 1542 responses to the online questionnaire in 2019 and was prepared for clustering with JMP and ANFIS implementation with MATLAB version R2022b. The ANFIS system was first presented by Jang [59], and its core is a fuzzy inference system that greatly simplifies the establishment of logical judgment [60]. K-means clustering is a simple unsupervised machine learning algorithm that is widely used for clustering tasks. The clustering results provided inputs for the ANFIS fuzzy system, which was used to predict the type of investment based on customer demographic groups. The fuzzy rules were generated based on input membership functions, and both system-generated and manually created rules were considered in this study. Data were collected through the Portfolio Investment questionnaire, and a subset of the data was used for clustering and ANFIS implementation. The proposed model provides a recommender system for suggesting investment types to potential investors based on investment type ratings.

4 The Proposed IRS Model

As previously stated, this research introduces a recommender system model based on ANFIS to provide investment recommendations. The model employs customer demographic data to cluster investment types and recommend them to potential investors. The model comprises four stages. The first stage involves collecting, storing, and initially processing data. In the second stage, machine learning techniques are applied to determine investment type clusters and customer groups based on demographic data. In the third stage, the ANFIS system is implemented. The fourth stage is to deliver suitable investment recommendations to potential investors through specific applications. After receiving the recommendations, customers are asked to complete a survey form to provide feedback on the system's performance.

The collected feedback is then used to improve the system recommendations in a repeated cycle. The proposed model consists of three primary phases: The first phase involves data collection, which has two layers—data acquisition and processing and data storage. The second phase is related to data analysis, comprising two functions—using machine learning techniques for data clustering and classification and ANFIS deployment. The third phase is the decision phase, where the recommended investment type is provided to the customer, and their feedback is received. Customer feedback helps identify any errors in the system, which are then corrected in the second phase.

5 Experiments and Results

This section comprises three parts: data clustering of investment types based on potential investor's responses, demographic data description for the ANFIS, and ANFIS inputs and outputs.

5.1 Clustering Investment Type Data

To implement the proposed IRS, we first clustered the data related to the investment types or products used by potential investors. We used the JMP software to cluster data based on four questions related to investment type. The investment type data involved various investment products, such as listed stock mutual funds, voluntary pension funds, government securities/bonds, and other financial products. We coded the answers given by potential investors and converted them into numerical data for analysis in MATLAB and JMP software. We prepared the data for clustering by using the K-means technique in JMP, resulting in the clustering of investment types into three clusters. All rows totaled 2038, with the first cluster containing 592 rows, the second cluster containing 406 rows, and the third cluster containing 340 rows. The K-means technique assigned data points to the nearest cluster center, and the reassignment process was repeated four times until the data points remained in their cluster. In case of adding new data in real-time, the K-means technique is recalculated, and the clusters are updated accordingly.

5.2 Demographic ANFIS

To gain insight into the relationship between demographics and attitudes toward finance, we asked six questions (inputs 1–6) related to demographic information. The responses aimed to help us better understand respondent's decision-making about investment types/products.

5.3 Demographic ANFIS Inputs and Output

The demographic ANFIS system for potential investors considered six inputs based on the following questions: (1) gender, (2) age, (3) location, (4) education, (5) job, and (6) income, and one output related to investment type clusters. Each input had specific membership functions (MFs). Input 1 (gender) had two MFs: "male" for MF1 and "female" for MF2. Input 2 (age) had three MFs: "15–34 years old" for MF1, "35–54 years old" for MF2, and "55–79 years old" for MF3. Input 3 (location) had two MFs: "Budapest" for MF1 and "other location" for MF2. Input 4 (education) had four MFs: "college or university economics" for MF1, "college or university non-economics" for MF2, "postgraduate" for MF3, and "other" for MF4. Input 5 (job) had nine MFs: "employee middle management" for MF1, "small medium business" for MF2, "graduate freelance" for MF3, "employed lower manager" for MF4, "subordinate intellectual worker" for MF5, "skilled worker" for MF6, "employed senior management" for MF7, "micro or self-employed" for MF8, and "other" for MF9. Input 6 (income) had three MFs: "under 200,000 HUF" for MF1, "200,000–349,999 HUF" for MF2, and "above 350,000 HUF" for MF3. The output was defined as one output related to investment type clusters, which included three clusters.

6 Implementation of DemographicANFIS

The ANFIS system operates based on "IF_THEN" rules using input membership functions (MFs). The ANFIS system architecture comprises three main parts:

Fuzzy Rules: ANFIS uses fuzzy rules to map input variables to output variables. These rules are typically expressed in the form of "IF-THEN" statements, where the antecedent (IF part) contains a fuzzy set defined by the input membership functions, and the consequent (THEN part) contains a linear or nonlinear function of the input variables.

Membership Functions (MFs): ANFIS uses MFs to map input variables to fuzzy sets. A fuzzy set is a set of values that are assigned a degree of membership between 0 and 1, representing the degree to which the input variable belongs to that fuzzy set. ANFIS typically uses Gaussian or triangular-shaped MFs.

Adaptive Network: The third part of the ANFIS system is the adaptive network, which is responsible for tuning the parameters of the fuzzy rules and MFs to optimize the mapping from input variables to output variables. The adaptive network is typically implemented as a feedforward neural network with a hybrid learning algorithm that combines gradient descent and least squares methods. The output of the ANFIS system is generated by combining the output of all the fuzzy rules using weighted averaging (Fig. 6).

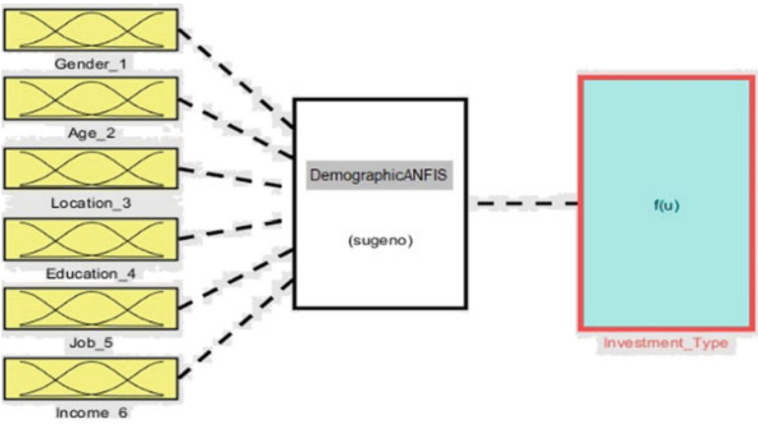


Fig. 6 Designing DemographicANFIS

Figure 7 illustrates a sample of the membership functions used in DemographicANFIS, which consist of a certain number of MFs. The type of MF used is gaussmfs, and the membership functions can be customized. The output MF type is kept constant. The model was trained using 1542 pairs of data. The implication method used is min, and the aggregation method used is max. In aggregation, all fuzzy sets that represent the outputs of each rule are combined into a single set, which is performed only once for each output variable before the final defuzzification stage.

The data were loaded for the next steps of data training and testing. To train the data, a new FIS was created using grid partition and a hybrid optimization method with an error tolerance of 0 and 3 epochs. The DemographicANFIS was generated as a new FIS, with 6 inputs for demographic groups and 1 output for investment types. During the training procedure, gaussmfs were used for each input. The FIS was trained using a hybrid method with 3 epochs, resulting in an error of 0.86685 after epoch 3. ANFIS training was then initiated, with a designated epoch number of 2. The training was completed at epoch 2, with a minimal training RMSE of 0.86683.

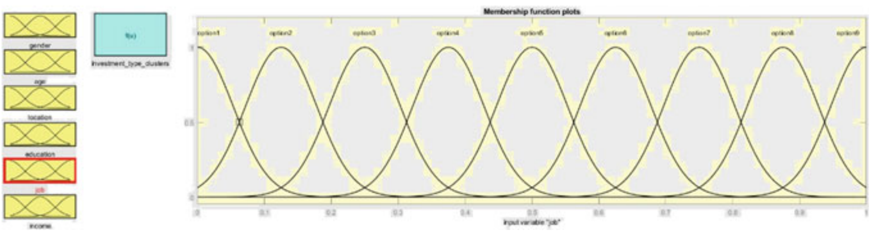


Fig. 7 Sample of membership functions in DemographicANFIS

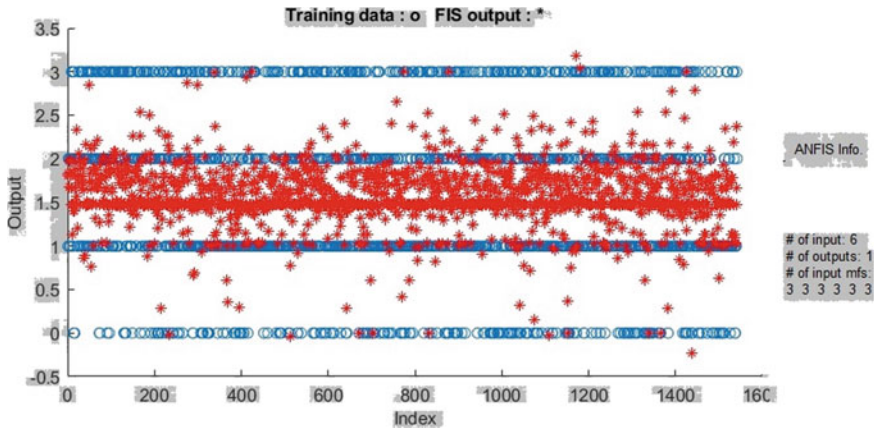


Fig. 8 Testing DemographicANFIS

Start training ANFIS ...

1. 0.866845

2. 0.86683

Designated epoch number reached. ANFIS training completed at epoch 2.

Minimal training RMSE = 0.86683

>>

The tested DemographicANFIS is depicted in Fig. 8, with an average testing error of 0.86683. Figure 8 displays a portion of the rule viewer, demonstrating the open system of the DemographicANFIS. With 1296 rules and 101 plot points, these rules can be added, modified, or removed based on expert insights and investor feedback, making this feature highly valuable for recommender systems. The proposed IRS structure of the DemographicANFIS is presented in Fig. 9.

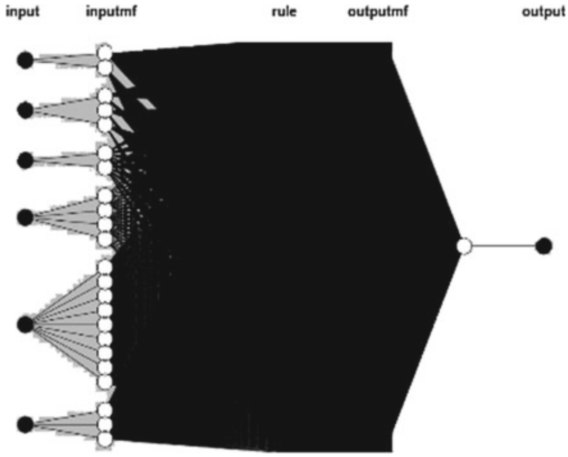


Fig. 9 Structure of DemographicANFIS (IRS)

7 Conclusion

In this research, an automated recommender system based on demographic characteristics was proposed to provide investment type suggestions to investors. The system utilizes a new intelligent approach using ANFIS that works with six demographic factors, even with incomplete data. The proposed model utilizes fuzzy neural inference and selection of investment types to provide advice and support the investor's decision-making. The model comprises seven group agents for the implementation of the IRS. The ANFIS system takes six demographic factors as inputs and provides a factor as output that corresponds to three clusters of investment types. The system has 2947 nodes and 1365 parameters, including 1296 linear parameters and 69 nonlinear parameters. The system was trained using 1542 training data pairs, and 1296 fuzzy rules were created. Based on the findings, it was concluded that the respondents of each cluster had specific features. However, the proposed system has limitations, including only considering demographic data of potential investors with six variables as inputs. Future research can consider additional characteristics of potential investors as inputs to the system. Expert opinions and feedback from investors can also be utilized to generate new fuzzy rules that were not produced by the system. Overall, the proposed system is a reliable recommendation system for investment products based on demographic characteristics.

References

1. Paranjape-Voditel P, Umesh D (2013) A stock market portfolio recommender system based on association rule mining. Appl Soft Comput 13(2):1055–1063
2. Kanaujia PKM, Manjusha P, Siddharth SR (2017) A framework for development of recommender system for financial data analysis. Int J Inform Eng Electron Bus 9(5):18–27

3. Hernández E, Sittón I, Rodríguez S, Gil AB, García RJ (2019) An investment recommender multi-agent system in financial technology. In: Graña M, López-Guede JM, Etxaniz O, Herrero Á, Sáez JA, Quintián H, Corchado E (eds) International joint conference SOCO'18-CISIS'18-ICEUTE'18, vol 771. pp 3–10
4. Tejeda-Lorente Á, Bernabé-Moreno J, Herce-Zelaya J, Porcel C, Herrera-Viedma E (2019) A risk-aware fuzzy linguistic knowledge-based recommender system for hedge funds. Proc Comput Sci 162:916
5. Faridniya A, Faridnia M (2019) Providing a model for allocating resources and choosing investment type using Data Envelopment Analysis (DEA) (Case Study: Social Security Organization). J Adv Pharmacy Educ Res 9(S2):112–124. https://japer.in/storage/models/article/et0pIWClvo41b1Rk0kK0g25dwiwg85RgsRsGFGDgP80KldRAN33ipHhEd1Rc/providing-a-model-for-allocating-resources-and-choosing-investment-type-using-data-envelopment-ana.pdf
6. Tarnowska K, Ras ZW, Daniel L (2020) Recommender system for improving customer loyalty. Springer International Publishing. https://doi.org/10.1007/978-3-030-13438-9
7. Sulistiyo H, Mahpudin E (2020) Demographic analysis for the selection of an investment type for amateur golfers. In: Advances in business, management, and entrepreneurship. CRC Press
8. Gottschlich J, Hinz O (2014) A decision support system for stock investment recommendations using collective wisdom. Decision Support Syst 59(1):52–62. Scopus. https://doi.org/10.1016/j.dss.2013.10.005
9. Salge TO, Kohli R, Barrett M (2015) Investing in information systems: on the behavioral and institutional search mechanisms underpinning hospital's is investment decisions. MIS Quart: Managem Inform Syst 39(1):61–89. Scopus. https://doi.org/10.25300/MISQ/2015/39.1.04
10. Zhou D, Yu Y, Wang Q, Zha D (2019) Effects of a generalized dual-credit system on green technology investments and pricing decisions in a supply chain. J Environ Managem 247:269–280. Scopus. https://doi.org/10.1016/j.jenvman.2019.06.058
11. Starita S, Scaparra MP (2016) Optimizing dynamic investment decisions for railway systems protection. Europ J Operational Res 248(2):543–557. Scopus. https://doi.org/10.1016/j.ejor.2015.07.025
12. Ullah F, Sepasgozar SME (2020) Key factors influencing purchase or rent decisions in smart real estate investments: a system dynamics approach using online forum thread data. Sustainability (Switzerland) 12(11). Scopus. https://doi.org/10.3390/su12114382
13. Kovačić D, Usenik J, Bogataj M (2017) Optimal decisions on investments in urban energy cogeneration plants—extended MRP and fuzzy approach to the stochastic systems. Int J Prod Econom 183:583–595. Scopus. https://doi.org/10.1016/j.ijpe.2016.02.016
14. Del Giudice V, De Paola P, Torrieri F, Nijkamp PJ, Shapira A (2019) Real estate investment choices and decision support systems. Sustainability (Switzerland), 11(11). Scopus. https://doi.org/10.3390/su11113110
15. Geressu RT, Harou JJ (2015) Screening reservoir systems by considering the efficient trade-offs—informing infrastructure investment decisions on the Blue Nile. Environ Res Lett 10(12). Scopus. https://doi.org/10.1088/1748-9326/10/12/125008
16. Yan Y, Hong L, He X, Ouyang M, Peeta S, Chen X (2017) Pre-disaster investment decisions for strengthening the Chinese railway system under earthquakes. Transp Res Part E: Logistics and Transport Rev 105:39–59. Scopus. https://doi.org/10.1016/j.tre.2017.07.001
17. Naranjo R, Santos M (2019) A fuzzy decision system for money investment in stock markets based on fuzzy candlesticks pattern recognition. Expert Syst with Appl 133:34–48. Scopus. https://doi.org/10.1016/j.eswa.2019.05.012
18. Fang S, Zhou P, Dinçer H, Yüksel S (2021) Assessment of safety management system on energy investment risk using house of quality based on hybrid stochastic interval-valued intuitionistic fuzzy decision-making approach. Safety Sci 141. Scopus. https://doi.org/10.1016/j.ssci.2021.105333
19. Babaei G, Bamdad S (2020) A multi-objective instance-based decision support system for investment recommendation in peer-to-peer lending. Expert Syst with Appl 150. Scopus. https://doi.org/10.1016/j.eswa.2020.113278

20. Lakhno V, Malyukov V, Gerasymchuk N, Shtuler I (2017) Development of the decision making support system to control a procedure of financial investment. Eastern-Europ J Enterprise Technol 6(3–90):35–41. Scopus. https://doi.org/10.15587/1729-4061.2017.119259
21. Teotónio I, Cabral M, Cruz CO, Silva CM (2020) Decision support system for green roofs investments in residential buildings. J Cleaner Prod 249. Scopus. https://doi.org/10.1016/j.jclepro.2019.119365
22. Mo J-L, Schleich J, Zhu L, Fan Y (2015) Delaying the introduction of emissions trading systems-implications for power plant investment and operation from a multi-stage decision model. Energy Econ 52:255–264. Scopus. https://doi.org/10.1016/j.eneco.2015.11.009
23. Kamari A, Jensen S, Christensen ML, Petersen S, Kirkegaard PH (2018) A hybrid decision support system for generation of holistic renovation scenarios-cases of energy consumption, investment cost, and thermal indoor comfort. Sustain (Switzerland) 10(4). Scopus. https://doi.org/10.3390/su10041255
24. von Appen J, Braun M (2018) Interdependencies between self-sufficiency preferences, techno-economic drivers for investment decisions and grid integration of residential PV storage systems. Appl Energy 229:1140–1151. Scopus. https://doi.org/10.1016/j.apenergy.2018.08.003
25. Renna P (2017) A decision investment model to design manufacturing systems based on a genetic algorithm and Monte-Carlo simulation. Int J Comput Integr Manuf 30(6):590–605. Scopus. https://doi.org/10.1080/0951192X.2016.1187299
26. Flora M, Vargiolu T (2020) Price dynamics in the European Union Emissions trading system and evaluation of its ability to boost emission-related investment decisions. Europ J Operat Res 280(1):383–394. Scopus. https://doi.org/10.1016/j.ejor.2019.07.026
27. Ali MN, Hameedi KS, Almagtome AH (2019) Does sustainability reporting via accounting information system influence investment decisions in Iraq? Int J Innov Creativity and Change 9(9):294–312. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85078994416&partnerID=40&md5=fed833934de74819c1b35daf8382616d
28. Kafuku JM, Saman MZM, Yusof SM, Sharif S, Zakuan N (2015) Investment decision issues from remanufacturing system perspective: literature review and further research. Proc CIRP 26:589–594. Scopus. https://doi.org/10.1016/j.procir.2014.07.043
29. Keding C, Meissner P (2021) Managerial overreliance on AI-augmented decision-making processes: how the use of AI-based advisory systems shapes choice behavior in R&D investment decisions. Technol Forecasting and Soc Change 171. Scopus. https://doi.org/10.1016/j.techfore.2021.120970
30. Jankova Z, Jana DK, Dostal P (2021) Investment decision support based on interval type-2 fuzzy expert system. Eng Econom 32(2):118–129. Scopus. https://doi.org/10.5755/j01.ee.32.2.24884
31. Ribas JR, da Silva Rocha M (2015) A decision support system for prioritizing investments in an energy efficiency program in favelas in the city of Rio de Janeiro. J Multi-Criteria Decis Anal 22(1–2):89–99. Scopus. https://doi.org/10.1002/mcda.1524
32. Akhmetov B, Balgabayeva L, Lakhno V, Malyukov V, Alenova R, Tashimova A (2019) Mobile platform for decision support system during mutual continuous investment in technology for smart city. vol 199. Springer International Publishing, pp 742. Scopus. https://doi.org/10.1007/978-3-030-12072-6_59
33. Quitoras MR, Cabrera P, Campana PE, Rowley P, Crawford C (2021) Towards robust investment decisions and policies in integrated energy systems planning: evaluating trade-offs and risk hedging strategies for remote communities. Energy Convers Managem 229. Scopus. https://doi.org/10.1016/j.enconman.2020.113748
34. Akhmetov B, Lakhno V, Malyukov V, Sarsimbayeva S, Zhumadilova M, Kartbayev T (2019) Decision support system about investments in smart city in conditions of incomplete information. Int J Civil Eng Technol 10(2):661–670. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063560122&partnerID=40&md5=e5ef10afb616fec417a3e09be9bec935
35. Bruaset S, Sægrov S, Ugarelli R (2018) Performance-based modelling of long-term deterioration to support rehabilitation and investment decisions in drinking water distribution systems. Urban Water J 15(1):46–52. Scopus. https://doi.org/10.1080/1573062X.2017.1395894

36. Li C-B, Yuan J-H, Gao P (2016) Risk decision-making based on Mahalanobis-Taguchi system and grey cumulative prospect theory for enterprise information investment. Intell Decis Technol 10(1):49–58. Scopus. https://doi.org/10.3233/IDT-150236
37. Akhmetov B, Kydyralina L, Lakhno V, Mohylnyi G, Akhmetova J, Tashimova A (2018) Model for a computer decision support system on mutual investment in the cybersecurity of educational institutions. Int J Mech Eng Technol 9(10):1114–1122. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85056288483&partnerID=40&md5=c795e0e5bd79f49ab2af5436f511dd2e
38. Cano EL, Moguerza JM, Ermolieva T, Yermoliev Y (2017) A strategic decision support system framework for energy-efficient technology investments. TOP 25(2):249–270. Scopus. https://doi.org/10.1007/s11750-016-0429-9
39. Al-Augby S, Majewski S, Nermend K, Majewska A (2016) Proposed investment decision support system for stock exchange using text mining method. In: Al-Sadiq international conference on multidisciplinary in IT and communication techniques science and applications, AIC-MITCSA 2016, pp 93–98. Scopus. https://doi.org/10.1109/AIC-MITCSA.2016.7759917
40. Cabrera-Paniagua D, Rubilar-Torrealba R (2021) A novel artificial autonomous system for supporting investment decisions using a big five model approach. Eng Appl Artif Intell 98. Scopus. https://doi.org/10.1016/j.engappai.2020.104107
41. Tao Z, Moncada JA, Poncelet K, Delarue E (2021) Review and analysis of investment decision making algorithms in long-term agent-based electric power system simulation models. Renew Sustain Energy Rev 136. Scopus. https://doi.org/10.1016/j.rser.2020.110405
42. Khalatur S, Khaminich S, Budko O, Dubovych O, Karamushka O (2020) Multiple system of innovation-investment decisions adoption with synergetic approach usage. Entrepreneurship and Sustain Issues 7(4):2745–2763. Scopus. https://doi.org/10.9770/jesi.2020.7.4(12)
43. Papapostolou CM, Kondili EM, Tzanes G (2018) Optimisation of water supply systems in the water—energy nexus: Model development and implementation to support decision making in investment planning, vol 43. Elsevier B.V., pp 1218. Scopus. https://doi.org/10.1016/B978-0-444-64235-6.50211-4
44. Siejka M (2017) The role of spatial information systems in decision-making processes regarding investment site selection. Real Estate Managem Valuation 25(3):62–72. Scopus. https://doi.org/10.1515/remav-2017-0023
45. Hu H, Zhou W (2014) A decision support system for joint emission reduction investment and pricing decisions with carbon emission trade. Int J Multimedia and Ubiquitous Eng 9(9):371–380. Scopus. https://doi.org/10.14257/ijmue.2014.9.9.37
46. Rühr A, Streich D, Berger B, Hess T (2019) A classification of decision automation and delegation in digital investment management systems. In: Proceedings of the annual Hawaii international conference on system sciences, 2019-January, pp 1435–1444. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85099474028&partnerID=40&md5=4b540866ec7ddd81afed3d00d76a58dc
47. Ortner J, Velthuis L, Wollscheid D (2017) Incentive systems for risky investment decisions under unknown preferences. Managem Account Res 36:43–50. Scopus. https://doi.org/10.1016/j.mar.2016.09.001
48. Li L, Cao X, Zhang S (2022) Shared energy storage system for prosumers in a community: investment decision, economic operation, and benefits allocation under a cost-effective way. J Energy Storage 50. Scopus. https://doi.org/10.1016/j.est.2022.104710
49. Sun J, Wang H, Chen J (2020) Decision-making of port enterprise safety investment based on system dynamics. Processes 8(10):1–17. Scopus. https://doi.org/10.3390/pr8101235
50. Xue J, Ye J, Tao Q, Wang D (2019) Multi-scenarios based operation mode and investment decision of source-storage-load system in business park. Dianli Zidonghua Shebei/Electric Power Autom Equipm 39(2):78–83 and 92. Scopus. https://doi.org/10.16081/j.issn.1006-6047.2019.02.012
51. Thomas EB, Dolecheck KA, Mark TB, Eastwood CR, Dela Rue BT, Bewley JM (2019) A decision-support tool for investment analysis of automated oestrus detection technologies in a seasonal dairy production system. Animal Prod Sci 59(12):2280–2287. Scopus. https://doi.org/10.1071/AN17730

52. Mutanov G, Milosz M, Saxenbayeva Z, Kozhanova A (2018) Investments decision making on the basis of system dynamics, vol 769. Springer, pp 303. Scopus. https://doi.org/10.1007/978-3-319-76081-0_25
53. Luo W (2020) Application of improved clustering algorithm in investment recommendation in embedded system. Microprocessors and Microsyst 75. Scopus. https://doi.org/10.1016/j.micpro.2020.103066
54. Wei Q, Li S, Gou X, Huo B (2019) Joint optimal decision of the shared distribution system through revenue-sharing and cooperative investment contracts. Indus Managem Data Syst 119(3):578–612. Scopus. https://doi.org/10.1108/IMDS-07-2018-0285
55. Ren K, Malik A (2019) Investment recommendation system for low-liquidity online peer to peer lending (P2PL) marketplaces. In: WSDM 2019—proceedings of the 12th ACM international conference on web search and data mining, pp 510–518. Scopus. https://doi.org/10.1145/3289600.3290959
56. Kozlova M, Collan M, Luukka P (2018) New investment decision-making tool that combines a fuzzy inference system with real option analysis. Fuzzy Econ Rev 23(1):63–92. Scopus. https://doi.org/10.25102/fer.2018.01.04
57. Niu D, Song Z, Wang M, Xiao X (2017) Improved TOPSIS method for power distribution network investment decision-making based on benefit evaluation indicator system. Int J Energy Sector Managem 11(4):595–608. Scopus. https://doi.org/10.1108/IJESM-05-2017-0005
58. Scaparra MP, Starita S, Sterle C (2015) In: Optimizing investment decisions for railway systems protection, vol 27. Springer, Netherlands, Scopus, pp 233. https://doi.org/10.1007/978-3-319-04426-2_11
59. Jang JR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man and Cybernet 23(3):665–685.https://doi.org/10.1109/21.256541
60. Asemi A, Asemi A (2014) Intelligent MCDM method for supplier selection under fuzzy environment. Int J Inf Sci Manag (IJISM) 12(2):33–40
61. Asemi A, Ko A (2021) A novel combined business recommender system model using customer investment service feedback. In: Proceeding of the 34th Bled eConference, June 27–30, 2021, Bled, Slovenia

I Am Bot the "Fish Finder": Detecting Malware that Targets Online Gaming Platform

Nicholas Ouellette, Yaser Baseri, and Barjinder Kaur

Abstract Malware in the gaming industry presents many forms of risk for users, such as phishing, trojans, malware, and network attacks. Few studies have been published on gaming industry malware. The ones identified were found to be primarily focused mainly on in-game cheat detection, such as cheat clients and aimbots. This paper leverages related research drawn from the broader field of cybersecurity, including, *URL-phishing detection and crytpojacking*. To detect URL phishing attacks data, we used eight filter, wrapper, and embedded-based feature selection and five machine learning techniques, i.e., AdaBoost, extra trees (ET), random forest (RF), decision tree (DT), and gradient boosting (GB) where highest accuracy, precision, recall, and F1_score are achieved with RF. We further scrutinize the feature selection and classifiers based on threshold that will help to provide an aggregate simplified recommendation to the user and alerting about the malicious URL. The outcome of whether the URL is benign or malicious can easily be seen on developed bot application named "*Fish Finder*." For allowing Discord users to protect themselves from future phishing attacks, we have shared the built application on github: https://github.com/Dinnerspy/Discord-Bot-Phishing-Detection.

Keywords Malware · Phishing · Machine learning · Online games · Cybersecurity

1 Introduction

As with other high-growth industries, malicious actors have increased their phishing and malware attacks on the gaming industry. Newzoo estimated that the total growth in the gaming industry for the year 2021 was 175.8 billion USD [1]. The key lead-

N. Ouellette (✉) · Y. Baseri · B. Kaur
University of New Brunswick, Fredericton E3B 9W4, Canada
e-mail: nicholas.ouellette@unb.ca

Y. Baseri
e-mail: yaser.baseri@unb.ca

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_21

ers in the gaming industry are currently Discord, Steam, Epic Games, and console platforms such as Xbox, Nintendo, and PlayStation. In addition, there are many new entrants in this space as new indie studios are starting up constantly, such as CCP Games, Jagex, and Psyonix Studios. As with any multi-billion-dollar industry, threat actors want to get their slice of the money; since the beginning of the COVID-19 pandemic, Kaspersky has estimated that there was more than a 50% increase in web-based gaming attacks [2]. The COVID-19 pandemic has primarily been viewed as a significant cause of this recent increase as there is a large increase in the number of people gaming from home [3]. No different than any other industry, end-users are always a target. Once an account is compromised, malicious actors have been known to change the password and sell the account on third-party websites to other users online. These accounts are valuable as they provide access to online games that had been purchased by the original account holder. This shows there is a significant chance of taking advantage of confidential user information. Due to these reasons, stopping phishing attacks in modern society is highly urgent even though they are challenge involved and risks of leakage of data. Blacklist and Whitelist approaches are the traditional methods to identify the phishing sites [4, 5]. A systematic review conducted in [6] describes various ML approaches to detect phishing attacks. The authors highlight the importance of using different feature types like URL and content-based features to effectively detect malicious In [7], the authors used 35 URL-based features and extracted term frequency and inverse document frequency (TF-IDF)-based features from content web pages to detect malicious. There proposed technique achieved an accuracy up to 98.25% with random forest.

After conducting research, we found that majority of existing research in the gaming industry is focused on cheating detection from a game server standpoint utilizing anomaly detection or by analyzing player performance [8–10]. Another common area of attack is within the PC gaming domain, where some users will attempt to install a “bootleg game” or install a software plugin that claims to provide the user with a competitive edge [11–13]. These attacks fall under the trojan category, as they typically install keyloggers or bitcoin miners [14] onto the victim’s computer.

Therefore, this study proposes using a subset of features which are selected using various feature selection techniques which will help ML to give better results. We also developed an application consisting of various bots that would alert online game users of legitimate or malicious phishing URL. Thus, to detect phishing website, following are the main contributions of this research:

- Firstly, we used different feature selection techniques to select best subset of features that helps in identifying malicious URLs.
- Secondly, we evaluated the performance of ML with selected features.
- Third, we developed a phishing detection application named ’Fish Finder’ that utilizes the functionality of best feature and ML combination which will timely alert online game users of any phishing attack.
- Finally, the Discord bot is available to the open-source community for continued improvement.

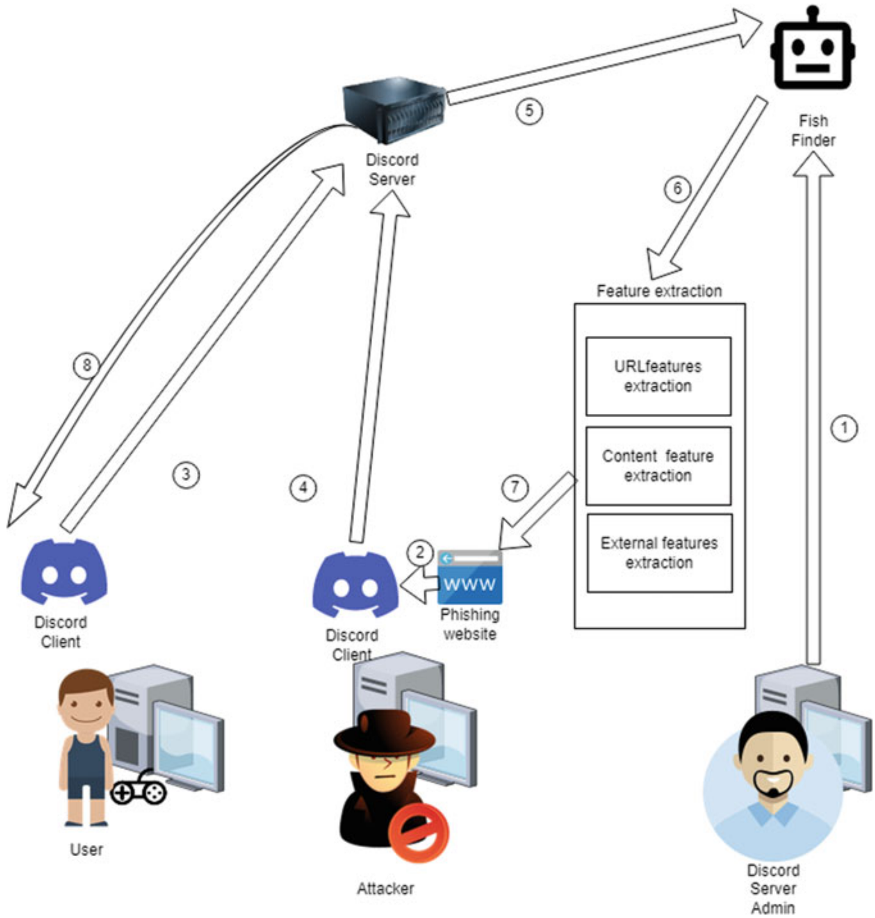


Fig. 1 Proposed framework using "Fish Finder" for phishing detection

The proposed flow is depicted in Fig. 1. Step 1: The Discord server admin starts up Fish Finder and it connects to Discord API in order to connect to Discord server. Step 2: Malicious actor creates a phishing URL based on a popular website. Step 3: Malicious actor logs into the Discord server and posts URL in text channel. Step 4: Average user connects to sever and reviews messages. Step 5: Fish Finder detects URL in post in text channel and sends to Fish Finder. Step 6: Fish Finder passes URL to the feature extraction. Step 7: Feature extraction is performed on the phishing URL and is passed back to them machine learning classifier. Step 8: The results of the prediction are sent back to the server, and the post is removed.

The rest of the paper is organized as follows: In Sect. 2, recent works in malware detection are discussed. In Sect. 3, the dataset details, feature selection, and methodologies followed are described. Experimental results are presented in Sect. 4. Finally, the conclusion and future directions are discussed in Sect. 5.

2 Related Work

Organizations and individuals are facing challenges related to phishing website due their similarity with legitimate website URLs. Phishers try to hack the credential information of the users by using different methods like forum postings, popping instant messages on the website, trying to retrieve information through URLs. The structure is so similar with real website that users get tricked and unintentionally causing severe economic damage to the intuitions. Researchers are working toward proposing and utilizing different ML approaches to timely determine whether the URL for opening the website is legitimate or malicious.

Bhardwaj et al. [15] developed a detection framework for new-age devices in order to mitigate phishing attacks. The authors noted that cybercriminals utilize new methods such as Python C&C servers with reverse tunneling applications such as NGROK. In their research, the authors set up a C&C (Command and Control server) along with NGROK as their phishing attack creation tools for their experiment. In order to combat these sophisticated attacks, the authors created a toolkit in Python that acts as a more secure DNS for end-users. The author's toolkit utilizes three phases for phishing detection. In the first phase, the toolkit filters traffic while utilizing malware scanning tools such as VirusTotal. The first phase also includes what the authors referred to as "phishing features," which tend to be traits that phishing websites have, such as long URLs, no Google index number or domain age. The second phase kicks in if a phishing attack gets through the first phase, and the user attempts to go to the website. Their toolkit will have a reliability pop-up appear based on some phishing features of the site for the end-user to decide to proceed or not, which ties into the third phase, which utilizes human interaction and learning about phishing which promotes human learning about phishing attacks.

Ripa et al. [16] conducted a look at the characteristics of phishing attacks and created a machine learning model that can detect phishing URLs, phishing emails, and phishing websites. The authors utilized approximately eleven thousand entries from the UCI Machine Learning Repository to train their model. The authors utilized six different machine learning training algorithms: Random forest, logistic regression, decision tree, Naive Bayes, KNN and XGBoost, and 70–30% dataset split to test performance. The authors found that their XGBoost classifier gave the highest accuracy and had the best performance of the six training methods for detecting phishing URLs. The authors found that their Naive Bayes classifier was the most accurate for detecting phishing emails, with an accuracy of about 95%. Finally, they found their random forest classifier performed the best for phishing website detection with an accuracy of about 96%.

Dutta [17] developed LURL, a machine learning algorithm utilizing a long short-term memory (LSTM) technique to determine between malicious and legitimate websites. To gather training data, the authors created a web crawler that got information on over 7900 URLs from the AlexaRank portal and the Phishtank dataset. However, utilizing the same LURL technique in different study [18, 19], the authors noted for the Phishtank dataset analyzing with LURL performed better overall with an F1 score of about 96. Further, from the crawler dataset, Le et al. [18] method performed better with an F1 of 95.6 compared to LURL's 94.8. The author expressed an interest to develop an unsupervised deep learning method with an extended study in future work.

Further, a study proposed by Rajab et al. [20] using new ranking features when experimented on UCI datasset attains an accuracy of 95.5% with C4.5 and JRIP classifiers, while analyzing their own collected dataset with 15 features authors in [21] found that RF reached highest phishing attack detection accuracy of 94.79%.

3 Materials and Methodology

3.1 Dataset

In this study, we have performed experiments on URL-based publicly available balanced dataset. This dataset contains 11,430 different legitimate and phishing website URLs with 87 extracted features which are divided into three classes: 56 URL-based, 24 content-based, and 7 external features. The authors collected legitimate web page URLs from Alexa and Yandex [22], while for URLs phishing data, both lists of URLs collected from Phishtank [23] and Openphish [24, 25]. We performed preporcessing on the dataset by removing null and duplicate values. We performed preporcessing on the dataset by removing null and duplicate values. In order to determine the best features and classifier model, a Python script was created to randomly split the dataset into 70% (training)/ 30% (testing) sets. A tenfold cross-validation is performed in order to capture an accurate performance.

3.2 Feature Selection

The main goal of feature selection step is to eliminate irrelevant or less contributing features. Selecting the best features can help reduce the dimensionality of the feature space, thereby saving us from the challenges associated with high-dimensional web data such as emails or websites. By selecting the best subset of overall features enhance the performance of the classifiers in providing correct predictions with less computation in detecting Phishing attacks. In this study, we have employed eight different filter, wrapper, and embedded-based feature selection techniques commonly used in phishing detection. The search procedures and specifications used by these techniques are presented in Table 1.

Table 1 Characteristics of features selection techniques

ID	Selection method	Search procedure	Specification
f_s1	CFS	Filter	Dependence
f_s2	mRmR	Embedded	Dependence
f_s3	Chi-squared	Filter	Transformation
f_s4	RFE	Wrapper	Subset selection
f_s5	Univariate	Embedded	Transformation
f_s6	Permutation	Wrapper	Dependence
f_s7	SFS	Wrapper	Subset selection
f_s8	SBS	Wrapper	Subset selection

- Correlation feature selection (CFS): It works by looking for features with high correlation with the target.
- Minimum redundancy maximum relevance (mRmR): It works by finding the minimal-optimal subset of features by removing irrelevant and useless features. This is done in order to find minimum amount of features that have the highest prediction ability.
- Chi-squared: It is designed to determine the dependency of the different features as compared to the classification variable utilizing a statistical model.
- Recursive feature elimination (RFE): It works by removing features from a dataset until the desired number of features is selected that are the most fit.
- Univariate feature selection: It utilizes univariate statistical tests to pick the best features for a given target variable.
- Permutation feature importance: It calculates feature importance by creating permutations of data subset and calculating model error in order to rank importance. All the features with a rank value of 0.01 or higher.
- Sequential feature selection (SFS): SFS starts with an empty set and fills it with features that do not decrease the value of the criterion.
- Sequential backward selection (SBS): SBS starts with the set full of features and removes them to pick the best subset so that it does not increase the value of the criterion.

By using different feature selection technique mentioned above, we have obtained subset of features depicted in Table 2.

3.3 Methodologies

This section discusses the approaches used to detect phishing attacks through URL. For detecting malicious URLs, we compared the performance of five classifiers which are explained below. Table 3 shows the hyper-parameter tuning of ML to achieve the results.

Table 2 Feature selection performed on different features used to detect phishing attacks

f_s	f_s1	f_s2	f_s3	f_s4	f_s5	f_s6	f_s7	f_s8
All features								
A_sID	A_s1	A_s1	A_s1	A_s1	A_s1	A_s1	A_s1	A_s1
	3, 6, 16, 21, 27, 51, 53557475798486	1, 10, 11, 21, 24, 26, 27, 29, 36, 41, 47, 51, 52, 55, 57, 58, 63, 75, 77, 78, 79, 80, 82, 83, 84, 85, 86, 87	3, 7, 21, 26, 27, 31, 34, 51, 56, 70, 78, 79, 83, 86, 87	2, 4, 5, 7, 11, 20, 21, 27, 40, 44, 45, 51, 57, 86, 87	1, 2, 3, 7, 10, 14, 21, 26, 51, 57, 58, 79, 83, 86, 87	5, 8, 10, 20, 21, 27, 47, 51, 71, 86, 87	5, 7, 11, 20, 21, 26, 36, 38, 51, 57, 58, 59, 63, 67, 70, 71, 83, 86, 87	2, 3, 5, 7, 11, 20, 21, 36, 47, 51, 57, 59, 63, 70, 71, 75, 83, 86, 87
$\#A_s$	13	28	15	15	15	11	19	19
URL-based features								
U_sID	U_s1	U_s2	U_s3	U_s4	U_s5	U_s6	U_s7	U_s8
	3, 7, 21, 27, 47, 51, 53, 55	1, 5, 10, 21, 26, 27, 29, 31, 36, 41, 47, 49, 51, 52, 55, 56	3, 7, 8, 10, 21, 26, 27, 31, 32, 34, 36, 51, 52, 55, 56	2, 4, 5, 6, 7, 10, 11, 14, 21, 26, 27, 45, 47, 50, 51	1, 2, 3, 4, 7, 10, 14, 21, 26, 27, 32, 34, 46, 47, 51	2, 4, 7, 10, 13, 14, 21, 26, 36, 40, 45, 47, 48, 49, 50, 51, 52, 55	2, 5, 7, 11, 14, 20, 21, 24, 25, 26, 27, 33, 34, 36, 50, 51, 52, 54, 55, 56	3, 4, 7, 10, 13, 14, 21, 25, 26, 27, 34, 36, 40, 48, 50, 51, 52, 55, 56
$\#U_s$	8	16	15	15	15	18	20	19
Content-based features								
C_s	C_s1	C_s2	C_s3	C_s4	C_s5	C_s6	C_s7	C_s8
	57, 59, 64, 73, 74, 79	57, 58, 59, 65, 66, 67, 70, 71, 73, 74, 75, 76, 77, 78, 79, 80	57, 58, 67, 68, 70, 71, 75, 78, 79, 80	57, 59, 61, 63, 67, 70, 71, 74, 79, 80	57, 59, 63, 67, 68, 70, 75, 78, 79, 80	57, 59, 63, 67, 70, 71, 79, 80	57, 58, 59, 61, 63, 64, 66, 67, 68, 70, 71, 73, 74, 75, 76, 77, 78, 79, 80	57, 58, 59, 60, 61, 63, 65, 66, 67, 68, 69, 70, 73, 74, 75, 76, 77, 78, 79, 80
$\#C_s$	6	16	10	10	10	8	19	20
External features								
E_sID	E_s1	E_s2	E_s3	E_s4	E_s5	E_s6	E_s7	E_s8
	84, 85, 86	81, 82, 83, 84, 85, 86, 87	81, 83, 85, 86, 87	81, 83, 85, 86, 87	81, 83, 85, 86, 87	83, 86, 87	81, 82, 83, 84, 85, 86, 87	81, 82, 83, 84, 85, 86, 87
$\#E_s$	3	7	5	5	5	3	7	7

Table 3 Hyper-parameters of machine learning algorithms to detect phishing attacks

ML approach	Parameters specification
Decision tree	Criterion = gini, min_samples_split = 2
Random forest	Number of trees =100, Max_depth = none
AdaBoost	n_estimators = 50, learning_rate = 1.0
Extra trees	Criterion = gini, n_estimators = 100
GB	Learning_rate = 0.1, min_samples_split = 2

- Decision tree (DT): It works by creating a tree structure in which leaves represent the class label and features are the branches that lead to the leaves.
- Random forest (RF): It works similar to decision trees, except during training time, many different decision trees are created. When performing classification, the prediction most represented in the different trees is selected.
- AdaBoost: It works by creating a forest of decision stumps based on the available features. This differs from a random forest in that it does not create all the trees at once but is based on the error of previous trees. In order to classify, it looks at the responses from all of the stumps.
- Extra Trees (ET): It works similar to random forest, except trees are not pruned. Also, each decision tree is based on the whole dataset.
- Gradient Boost (GB): It works similarly to AdaBoost, except trees are of a fixed size that can be larger than a stump. Also, like with AdaBoost, trees are scaled (given a more significant weight) based on importance.

4 Results

This section highlights the results obtained after analyzing using eight different feature selection techniques and five classifiers where Table 2 depicts the list of features selected by various techniques. However, all the ML approaches used for the experiment gave above 90% results for accuracy, precision, recall, and F1_score. But among them, RF and ET outperformed by providing above 95% results. When these subsets of features are fed into different classifier models results shown in Table 4, we noticed that with RF, we achieved highest accuracy of 95.28% and F1_Score of 95.33%, whereas with ET classifier, an accuracy of 95.15%, F1_score 95.18% with f_s2 feature selection technique to detect malicious URLs.

To leverage the solution provided by our proposed approach, we have developed a phishing detection Bot application named "Fish Finder" shown in Fig. 2 for Discord platform. This bot analyzes messages that are typed on the server looking for URLs within the chat messages. As threshold-based scheme proves beneficial for attack detection [26] thus by utilizing this scheme, we used the combination which provides above 95% results for further analysis. The analyses are performed by selected subset

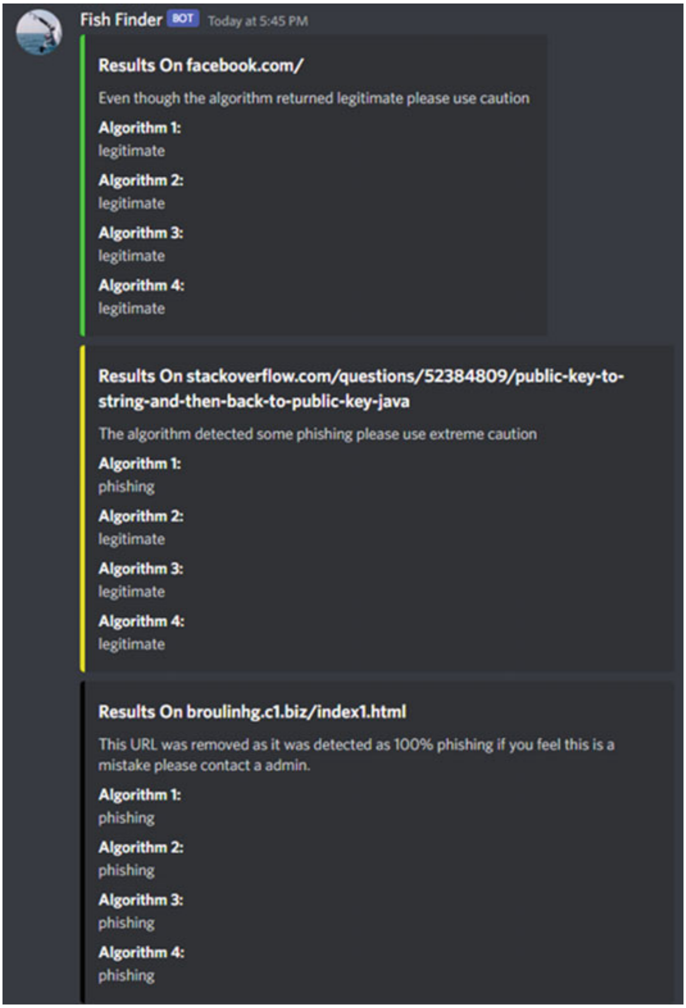


Fig. 2 Screenshot of bot application "Fish Finder" showing the status when user logs on to the server

of features and classifier combination models, i.e., "f_s2 with RF," "f_s4 with ET," "f_s5 with RF," and "f_s8 with RF" to detect phishing URLs and lower false positives. When found, the bot uses the selected algorithms to determine if the URL has a high probability of being a phishing URL. We have performed various analysis on subclasses of features, i.e., URL-based (56), content-based (24), and external features (7) using feature selection technique experiments and presented the results by the combination of best feature selection f_s2 and classifier (RF) which gave highest results as shown in Fig. 3.

Table 4 Results obtained using different feature selection techniques by classifiers

Approach	f_s1	f_s2	f_s3	f_s4	f_s5	f_s6	f_s7	f_s8
F-Score								
Decision tree	91.40	91.22	89.29	92.58	91.91	92.36	91.35	91.59
Random forest	92.35	**95.33**	92.08	**95.19**	**95.22**	93.68	94.79	**95.02**
AdaBoost	91.64	93.96	91.15	94.08	93.65	92.68	93.76	94.01
Extra Trees	91.26	**95.18**	90.94	**95.26**	**95.18**	93.24	94.56	**95.00**
Gradient boosting	92.90	94.99	91.84	94.69	94.80	93.51	94.71	94.65
Recall								
Decision tree	92.35	91.53	88.15	93.04	91.87	91.68	91.65	91.28
Random forest	92.97	**95.65**	92.35	**95.53**	**95.67**	93.96	**95.00**	**95.28**
AdaBoost	91.70	93.79	91.48	94.04	93.41	92.22	93.46	93.61
Extra trees	92.05	**95.25**	90.81	**95.25**	**95.45**	92.87	94.88	**95.03**
Gradient boosting	92.97	**95.18**	92.79	94.78	94.93	93.69	94.83	94.81
Precision								
Decision tree	90.48	90.92	90.4	92.13	91.94	93.06	91.04	91.90
Random forest	91.74	**95.01**	91.82	94.87	94.78	93.40	94.58	94.76
AdaBoost	91.59	94.14	90.82	94.13	93.88	93.15	94.05	94.41
Extra Trees	90.48	**95.11**	91.08	**95.28**	94.91	93.61	94.25	94.98
Gradient boosting	92.83	94.80	90.90	94.59	94.67	93.34	94.59	94.50
Accuracy								
Decision tree	91.26	91.13	89.36	92.50	91.86	92.37	91.26	91.56
Random forest	92.25	**95.28**	92.01	**95.15**	**95.17**	93.62	94.75	94.97
AdaBoost	91.58	93.93	91.06	94.05	93.62	92.67	93.73	94.00
Extra trees	91.12	**95.15**	90.90	**95.23**	**95.13**	93.22	94.51	94.97
Gradient boosting	92.85	94.95	91.70	94.65	94.76	93.46	94.67	94.61

Values were bolded if they were higher than or equal to 95%

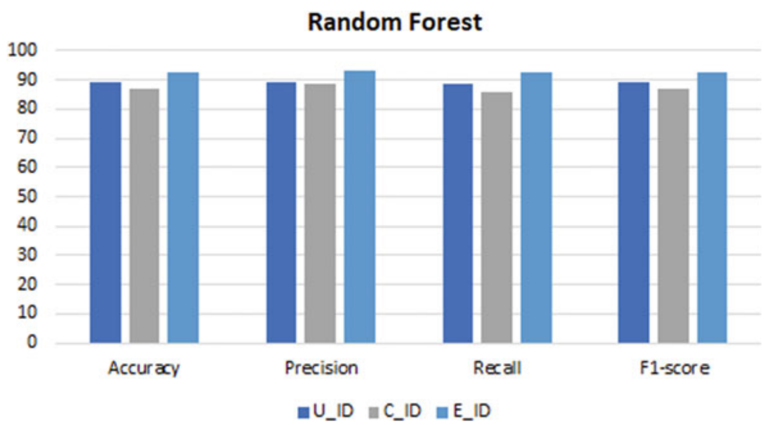


Fig. 3 Results presented obtained by using best feature and classifier combination

An aggregate recommendation is provided to the end-user that can be acted on, i.e.: remove the URL from the server, automatically or via manual admin intervention. The bot provides flexibility in order to either provide the results to the admin for a manual decision, or the admin can configure the bot to automatically remove the URL. Since four different machine learning models are being used, a level of aggregate confidence can be provided, whereby the automated removal can be configured to only remove URLs when all four are in agreement that the URL has a high probability of being phishing.

A developer Discord account was required from the Discord developer portal to create a Discord application to receive API access. From there, Discord py was utilized as the in-between tool for accessing the API for Discord and the machine learning models. The Discord bot was designed to analyze the chat from the channels. The Discord bot scans each message to find messages containing URLs. Once a message is found with a URL, the bot sends the URL to an adapted version of the data extractor created by Hannousse and Yahiouche [25]. The URL data extractor was modified to only collect the common URL features between f_s2," "f_s4," "f_s5," and "f_s8" thus collecting 41 features. The extracted features are split up into corresponding subsets based on the algorithms' requirements. From there, the algorithms predict the outcome determining whether the URL is phishing or legitimate. The bot adds a reaction to the message indicating the status of what it found:

- green for legitimate,
- yellow for 1/4 phishing detection,
- orange for 2/4 phishing detection,
- red for 3/4 phishing detection,
- X if there is an error reading the URL,
- Finally, the post is removed by the bot if all the algorithms affirm phishing attack.

All results of the bot are logged in a channel on the server showing the outcome of each URL. Figure 4 shows an example of bot being utilized with different URLs highlighting three different scenarios. It can be seen that of the three posts, only two showed up as the known phishing post was flagged as phishing and removed, whereas the other two were given status levels of green and yellow. In the Fish Finder logs, it can be noted that for the StackOverflow post, only one of the algorithms detected phishing which is why it received a status of yellow, and when looking at the phishing URL, all algorithms detected phishing and determined to remove the post

- Scenario 1–100% legitimate link with 4/4 pass: a Facebook page,
- Scenario 2–100% legitimate link with 3/4 pass: a StackOverflow post,
- Scenario 3—a Phishing website: a known phishing website found by the phishing detection

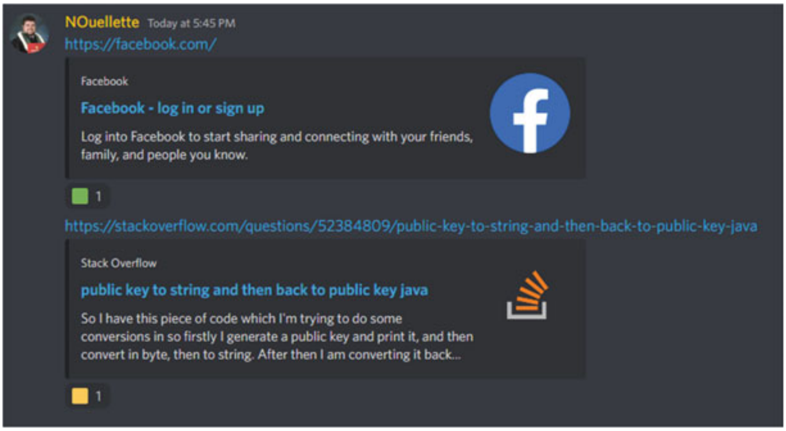


Fig. 4 Screenshot of the Discord text channel shows how the users see the results of three scenarios of the types of URLs that users could submit. The scenarios involved utilizing one Facebook link, which was legitimate, a StackOverflow post that was legitimate and a known phishing website. The Fish Finder Discord phishing detection bot updated the posts that were not detected as 100% phishing with emoji reactions that symbolize confidence in whether it is phishing or not (green for legitimate and yellow for 1/4 phishing detection). In the case of the third post, the bot removed it as it was detected as 100% phishing.

5 Conclusion and Discussion

With the gaming industry taking in $175 billion dollars annually [1], it is a target for cybersecurity threats. Kaspersky.com has noted there have been recent increases in attacks within the gaming industry. Due to lack of research on phishing attacks in online gaming, this study performed various experiments using combination of feature selection and ML approaches. Using this combination, we have analyzed that "mRmR" feature selection technique and RF classifier prove beneficial in providing highest results that will help in timely detection of phishing attacks. Further, we selected the best subset of feature and classifier combination and developed a bot application "Fish Finder" that will alert the online game player about legitimate or malicious URL.

There are a number of possible areas for future improvements, including performance optimization, simplification, and dataset enhancements.

- The current program developed can only process a single message at a time with messages queued for sequential processing. Larger Discord servers could run into latency issues if there were a high volume of messages generated in a short period of time. Parallel processing could be implemented to address this potential issue.
- In the future, a study could be conducted on the benefits of leveraging multiple machine learning algorithms to look for opportunities to potentially reduce the number used and improve the aggregate result. We will also work on different deep learning approaches.

- A final area of opportunity is to increase the size of the dataset in the ML models, by adding additional phishing and legitimate websites with URLs typically used on Discord chat servers. The current dataset is small and has limitations, as it was noted when testing the Discord bot that some of the machine learning algorithms would indicate a false phishing detection on a legitimate long URL (ex: Facebook photo URL).

6 Availability

The proposed ML Discord bot (Fish Finder) is available (under the LGPL-2.1 license) here: https://github.com/Dinnerspy/Discord-Bot-Phishing-Detection for people to install on their Discord server and/or contribute to future versions.

References

1. Kaspersky (2021) Analytical report on gaming-related cyberthreats in 2020–2021, May 2021. [Online]. Available: https://securelist.com/game-related-cyberthreats/103675/
2. Kaspersky (2021) Gaming-related web attacks increased by more than 50% in April, May 2021. [Online]. Available: https://www.kaspersky.com/about/press-releases/2020_gaming-related-web-attacks-increased-by-more-than-50-in-april
3. Vaas L Pandemic-bored attackers pummeled gaming industry. [Online]. Available: https://threatpost.com/attackers-gaming-industry/167183/
4. Srinivasa Rao R, Pais AR (2017) Detecting phishing websites using automation of human behavior. In: Proceedings of the 3rd ACM workshop on cyber-physical system security, pp 33–42
5. Almseidin M, Zuraiq AA, Al-Kasassbeh M, Alnidami N (2019) Phishing detection based on machine learning and feature selection methods
6. Dou Z, Khalil I, Khreishah A, Al-Fuqaha A, Guizani M (2017) Systematization of knowledge (sok): a systematic review of software-based web phishing detection. IEEE Commun Surveys Tutorials 19(4):2797–2819
7. Rao RS, Vaishnavi T, Pais AR (2020) Catchphish: detection of phishing websites by inspecting urls. J mbient Intell Humanized Comput 11(2):813–825
8. Witschel T, Wressnegger C (2020) Aim low, shoot high: evading aimbot detectors by mimicking user behavior. In: Proceedings of the 13th European workshop on systems security, ser. EuroSec '20. New York, NY, USA: association for computing machinery, pp 19–24. [Online]. Available: https://doi.org/10.1145/3380786.3391397
9. Qian X, Sifa R, Liu X, Ganguly S, Yadamsuren B, Klabjan D, Drachen A, Demediuk S (2022) Anomaly detection in player performances in multiplayer online battle arena games. In: Australasian computer science week 2022, ser. ACSW 2022. New York, NY, USA: Association for Computing Machinery, pp 23–30. [Online]. Available: https://doi.org/10.1145/3511616.3513095
10. Pinto JP, Pimenta A, Novais P (2021) Deep learning and multivariate time series for cheat detection in video games. Mach Learn 110(11):3037–3057. [Online]. Available: https://doi.org/10.1007/s10994-021-06055-x
11. Cyware (2020) Nitrohack: another malware turns discord client into a trojan: cyware hacker news. Jun 2020. [Online]. Available: https://cyware.com/news/nitrohack-another-malware-turns-discord-client-into-a-trojan-a67835b1/

12. Cyware (2020) Nintendo accounts hack: Hackers playing the wrong way: cyware hacker news. [Online]. Available: https://cyware.com/news/nintendo-accounts-hack-hackers-playing-the-wrong-way-6b8e45d5
13. Cyware (2020) How various malware families steal your gaming data: cyware hacker news. [Online]. Available: https://cyware.com/news/how-various-malware-families-steal-your-gaming-data-1f491259
14. Tekiner E, Acar A, Uluagac AS, Kirda E, Selcuk AA (2021) Sok: cryptojacking malware. In: IEEE European symposium on security and privacy (EuroS P), pp 120–139
15. Bhardwaj A, Al-Turjman, Sapra V, Kumar M, Stephan T (2021) Privacy-aware detection framework to mitigate new-age phishing attacks. Comput and Electr Eng 96:107546. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790621004912
16. Ripa SP, Islam F, Arifuzzaman M (2021) The emergence threat of phishing attack and the detection techniques using machine learning models. In: 2021 International conference on automation, control and mechatronics for industry 4.0 (ACMI), pp 1–6
17. Dutta AK (2021) Detecting phishing websites using machine learning technique. PLoS ONE, vol 16, no 10, pp 1–17. [Online]. Available: https://login.proxy.hil.unb.ca/login?; https://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=152954439&site=ehost-live&scope=site
18. Le H, Pham Q, Sahoo D, Hoi SCH (2018) Urlnet: Learning a URL representation with deep learning for malicious URL detection. CoRR, vol. abs/1802.03162, 2018. [Online]. Available: http://arxiv.org/abs/1802.03162
19. Hong J, Kim T, Liu J, Park N, Kim S-W (2020) Phishing URL detection with lexical features and blacklisted domains. Springer International Publishing Cham, pp 253–267. [Online]. Available: https://doi.org/10.1007/978-3-030-33432-1_12
20. Rajab KD (2017) New hybrid features selection method: a case study on websites phishing. Secur Commun Netw 2017
21. Stiawan D (2020) Phishing detection system using machine learning classifiers
22. Yandex Yandex.xml. [Online]. Available: https://yandex.com.tr/dev/xml/
23. phishtank, Join the fight against phishing [Online]. Available: https://www.phishtank.com/
24. OpenPhish, Phishing intelligence. [Online]. Available: https://openphish.com/
25. Hannousse A, Yahiouche S (2021) Towards benchmark datasets for machine learning based website phishing detection: an experimental study. Eng Appl Artif Intell 104:104347. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197621001950
26. David Akande T, Kaur B, Dadkhah S, Ghorbani AA (2022) Threshold based technique to detect anomalies using log files. In: 2022 7th international conference on machine learning technologies (ICMLT), pp 191–198
27. Namestnikova M (2021) Do cybercriminals play cyber games in quarantine? a look one year later. [Online]. Available: https://securelist.com/do-cybercriminals-play-cyber-games-in-quarantine-a-look-one-year-later/103031/
28. Muncaster P (2020) Stalker online breach: 1.3 million user records stolen. [Online]. Available: https://www.infosecurity-magazine.com/news/stalker-online-breach-13-m-user/
29. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830
30. AutoViML, Autoviml/featurewiz: Use advanced feature engineering strategies and select best features from your data set with a single line of code [Online]. Available: https://github.com/AutoViML/featurewiz
31. U. of Waikato, Weka 3: machine learning software in java. [Online]. Available: https://www.cs.waikato.ac.nz/~ml/weka/

The Application of Remote Sensing and GIS Tools in Mapping of Flood Risk Areas in the Souss Watershed Morocco

Jada El Kasri, Abdelaziz Lahmili, Ahmed Bouajaj, and Halima Soussi

Abstract The Souss watershed in the Souss-Massa region of Morocco is dominated by a semi-arid climate. The rainfall pattern is highly variable with frequent droughts and floods. The changing climatic conditions and the ongoing trends of overexploitation of water resources exacerbated the frequency and intensity of these climate hazards. The changing hazard characteristics have a significant impact on the communities who largely depend on agriculture, tourism and fishing activities. Detailed characterization of hazard risks is the first step for management of risks associated with climate extremes and optimal use water resources. This study presents a comprehensive methodology for mapping of flood risk areas to support decision-making in relation to flood hazard risk management, including flood prevention and preparedness. The methodology employs a multicriteria analysis in a GIS environment for identification of vulnerable flood-prone areas. The methodology involves combination of several thematic maps representing the most determining factors of flooding. The factors include geographic (elevation, slope, curvature, lithology) biophysical (land use, Normalized Difference Vegetation Index-NDVI) and hydrological (rainfall, stream density, stream power index, topographic wetness index) features. The weighted values of the parameters were used for mapping based on their relative importance for flood occurrence and severity. With the weighted values, the maps of the above parameters were overlaid and flood hazard maps were produced. The resulting maps were used to divide the watershed into five flood severity regions. The historical flood event data from the Souss-Massa Draa Hydraulic Basin Agency (ABHSMD) was used to validate the reliability of the flood hazard maps. The results showed that the parameters selected for mapping capture the variability of flood risks and could be used for planning and decision-making for both flood risk and water resources management.

J. E. Kasri (✉) · A. Lahmili · H. Soussi
Laboratory of Applied Geophysics, Geotechnics, Engineering and Environmental Geology (L3GIE), The Mohammadia School of Engineers, Mohammed V University, Rabat, Morocco
e-mail: jada.elkasri@gmail.com

A. Bouajaj
Laboratory of Engineering Sciences and Applications, National School of Applied Sciences, Abdelmalek Essaâdi University, Al-Hoceima, Morocco

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_22

Keywords Flood risks · Souss watershed · Weighted overlay · Geographical information systems (GIS)

1 Introduction

Climate change is increasing the frequency and intensity of irregular weather events, including erratic precipitation, variable in both temporal and spatial dimensions and increasing temperatures. These changes are leading to a variety of biophysical and socioeconomic problems. Localized high-intensity rainstorms, floods, droughts and heat waves could often damage infrastructure and cause loss of lives [1, 2]. The frequency of floods in some regions is increasing and is associated with heavy rainfall causing disasters in areas where livelihood activities are dominant [3]. Floods often damage the roads, rail networks, bridges, infrastructure of electricity and water distribution, and services such as telephone. The widespread damage to these infrastructure not only causes the huge damage to economic activities but also leads to significant economic cost to build back better, especially in areas where high levels of infrastructure and transportation facilities exist [4]. The flash floods are a common phenomenon in areas where high-intensity rainfall occurs within a shorter period of time and where topographic features favour rapid flow of water [5]. Further, flash floods are common in areas where a combination of meteorological, hydrological and human parameters that are not favourable for smooth flow and drainage of water [6]. Therefore, the adoption of a suitable flood prevention and mitigation measures depending on the flood characteristics is crucial.

The flood prevention and mitigation measures must take into account information on vulnerability to flooding and flood zones. In Souss watershed of Morocco, the intensity of these extreme events is becoming more frequent and alarming with climate change as it was reported in the previous study [7]. The extreme events of flooding and drought are aggravated both by the changes in meteorological variables (increase in temperatures and the decrease in rainfall) and by the inherent characteristics of the physical environment (high slope, impermeable rocks). Changes in the demography with more population, unsustainable land use, overexploitation of natural resources, deforestation are considered factors contributing to the frequent extreme floods [8]. In this study, the analysis focuses on the identification and mapping of flood-prone areas in order to help planning for appropriate mitigation and prevention measures. The area is considered flood risk if the probability of flooding event of a certain degree will occur in a specific area during certain period of time. The flood risk maps are often used for land use planning and selection of agricultural practices. The spatial flood maps are easy-to-read and accessible to decision-makers at various levels to identify risk areas and prioritize their flood prevention and mitigation measures [9]. Modelling techniques such as hydrological and hydraulic modelling and global hydrodynamics modelling are being widely used by many researchers [10–13]. Similarly, the weighted overlay technique has been applied in mapping in several studies that focused on potential water catchment areas mapping [14], landslide susceptibility mapping [15] and urban flood hazard detection [16].

2 Study Area

The Souss watershed is situated in the Souss-Massa region of Morocco and has an area of 16,200 Km^2. The region's climate is semi-arid, and highly variable rainfall pattern contributes to vulnerability; (see Fig. 1). Agriculture, fishing and tourism are the important livelihood activities of the region. The elevation of the watershed ranges from mean sea level to 4000 m above mean sea level. The high-altitude water catchments from north, east and south slope to westwards into the basin towards the Atlantic Ocean (Table 1).

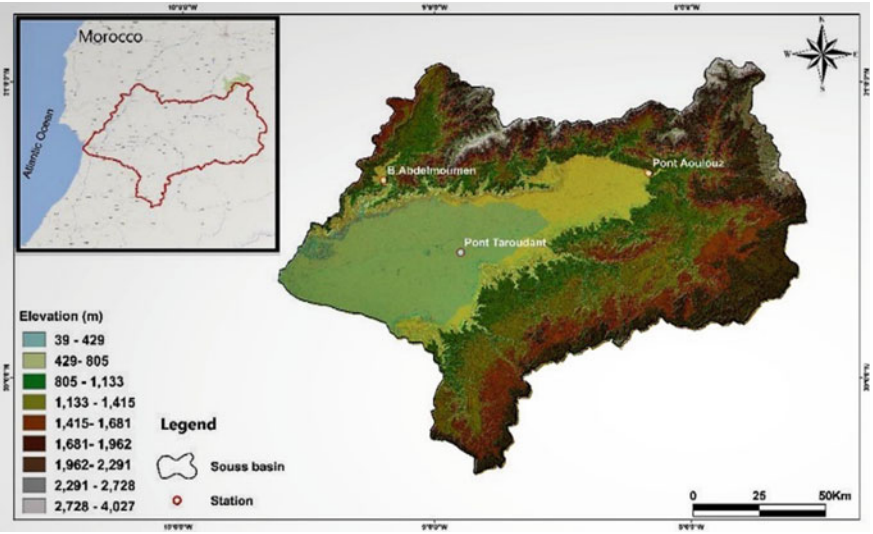


Fig. 1 Map of the study area

Table 1 Annual maximum and minimum temperatures in the main meteorological stations

Stations	Annual maximum temperature (°C)	Annual minimum temperature (°C)
PT	45.6	1.71
PA	43.2	3.65
BA	42 0.2	4.37

3 Materials and Methods

Flood risk mapping was carried following five major sequential steps: (1) collection of data and relevant parameters, (2) preparation of spatial datasets of the parameters, (3) assessment of flood risk by using the weighted overlay model, (4) preparation of flood risk maps and (5) validation of the maps using the data on flooded areas due to the historical flood events (Fig. 2). All figures are created by the authors, and no copyright is required (Fig. 3).

The mapping exercise used the weighted overlay technique and is considered as one of the most suitable multicriteria evaluation methods [17]. The method uses the geographical information system (GIS) for data management and for knowledge-based multicriteria analysis to combine value-added information with factual information [18]. The method is also used to develop multiple raster layers by giving weight to each raster layer depending on their importance [19]. Each of the raster

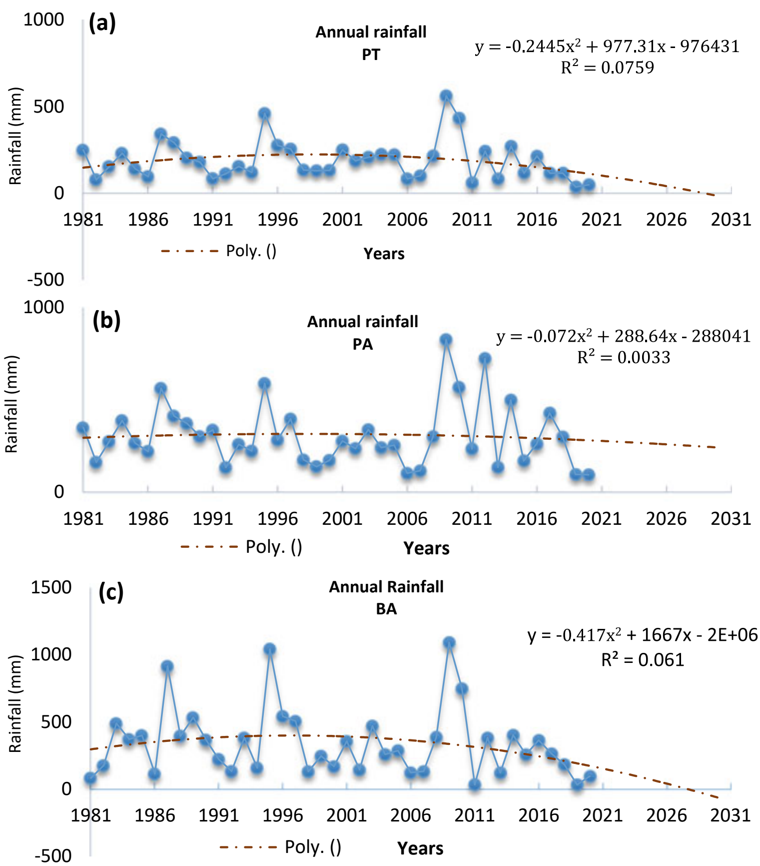


Fig. 2 Annual rainfall in the Souss watershed (three stations) over 40 years

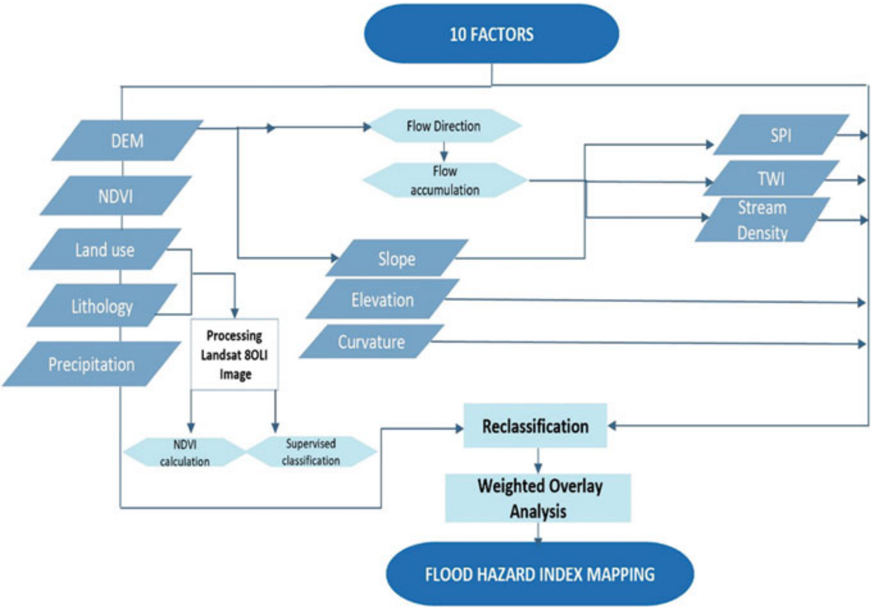


Fig. 3 Flowchart of the methodology

layers was assigned with weights based on their importance determined by experts' knowledge and opinion. The determinants of the flood risk maps can vary based on the biophysical and other characteristics of the watershed [20]. Thus, identification and characterization of these factors are necessary for flood modelling [21]. As indicated above, there are 10 factors identified for the Souss watershed for mapping.

According to a comprehensive literature review [2, 5, 25] and through an examination from a field survey, characteristics of the historical floods that occurred in the Souss watershed were analysed and most prominent factors responsible for floods occurrence were identified. These factors include watershed's geographical (slope, elevation, curvature, lithology), biophysical (NDVI, land use) and hydrological properties (stream power index (SPI), topographic moisture index (TWI), rainfall and stream density). Separate thematic layers were developed for each of the above listed factors. Digital Elevation Model (DEM) was used for development of thematic layers of elevation, slope and flow accumulation. The land use information is consolidated from the normalized difference vegetation index (NDVI) and mapped. Rainfall intensity is estimated from measured rainfall covering the area of the watershed.

3.1 Factors Conditioning Flooding

Slope. Flood occurrence is determined by the slope angle [22, 23] among several other factors. When the slope angle is high, the water infiltration rate is low and thus creating higher water velocity and flow downstream. The rapid flow of high volume of water downstream reaching the low-lying areas captures huge volume of water within a shorter period of time. The situation creates flooding in low-lying areas, and subsequently, these areas are often prone to flooding [24]. The slope map was created with eight intervals (Fig. 4).

Elevation. It is also one of the factors associated with flooding [25]. In general, the areas of the watershed in lower elevation are prone to flooding as excess water from higher elevation throughout the watershed accumulates rapidly in lower elevation areas [26] in both central and west central parts of the watershed. The elevation map was developed with nine intervals and presented in (Fig. 5).

Curvature is used to determine the flooding in the watershed. In general, curvature is categorized as concave (negative curvature), flat (zero curvature) and convex (positive curvature). The curvature affects the surface runoff and infiltration characteristics [26]. The map representing the curvature is classified into concave, convex and flat (no curvature) and presented in (Fig. 6). The map clearly shows that the watershed is covered largely by flat curvature as evidenced from the characteristics of the Souss plain.

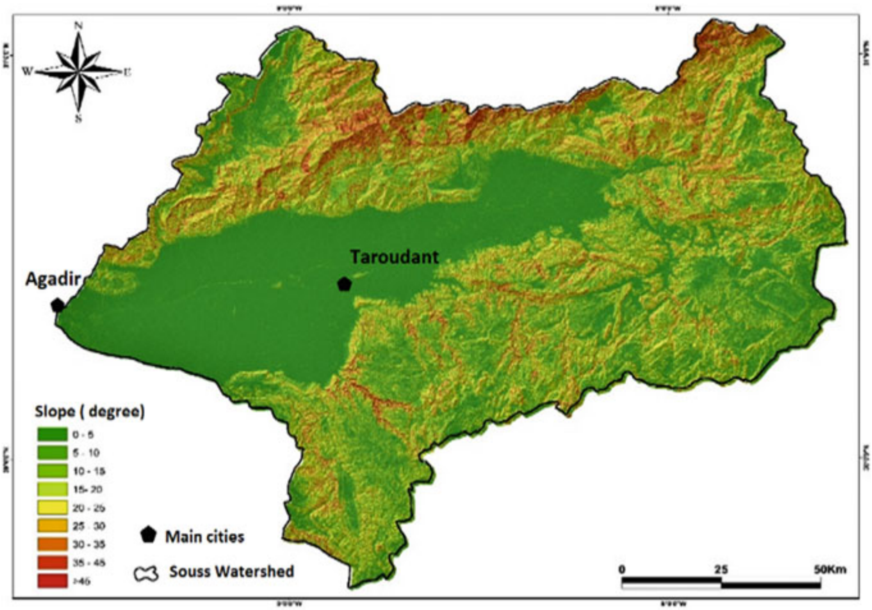


Fig. 4 Slope map in Souss watershed

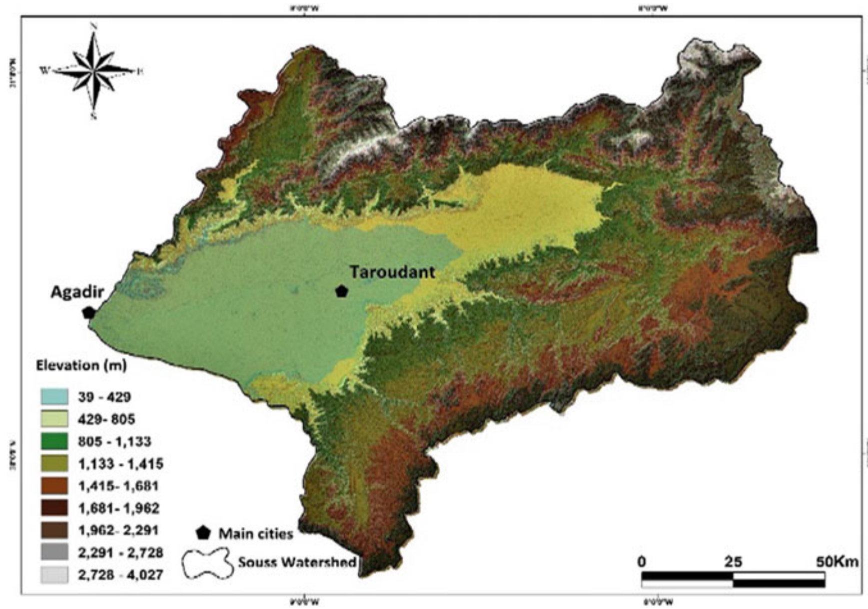


Fig. 5 Elevation map of the Souss watershed

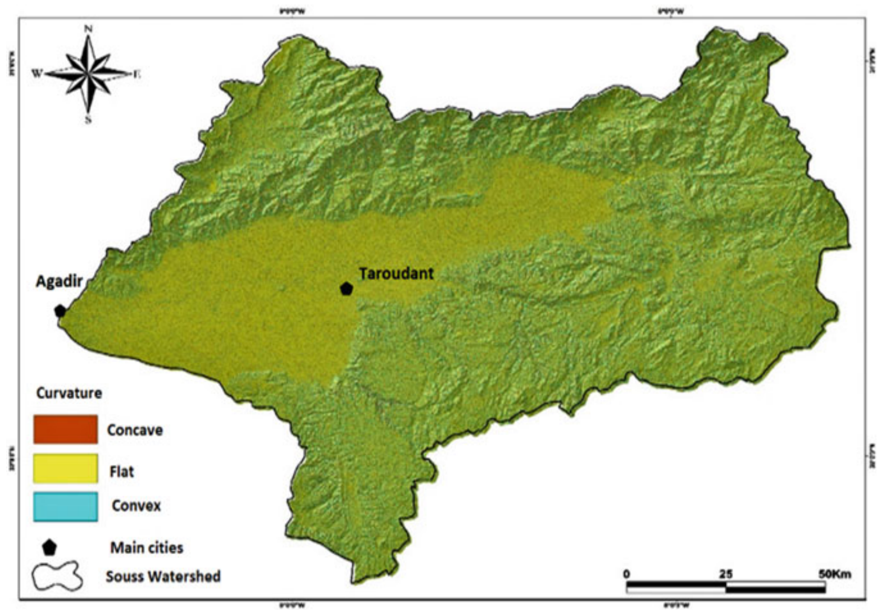


Fig. 6 Curvature map in Souss watershed

The Stream Power Index (SPI). The Stream Power Index (SPI) is used to characterize erosive power of the basin and relative flow rate in the watershed [27]. SPI is function of the soil moisture content and the power of floods to flow downstream within the watershed [26]. SPI reflects the abrasive power of floods. Higher the value of SPI means higher the flood power, whereas the lower the value of SPI means lower the flood power, but still there is a potential for flow accumulation in the watershed [28]. The SPI of the watershed was calculated with the following method:

$$\mathrm{SPI} = \mathrm{As}\tan \mathrm{g}\beta \tag{1}$$

where AS is the specific basin area and is the local slope gradient (in degrees). The SPI map consisted of three intervals (Fig. 7). As evidenced from the map, the high-elevation areas are dominated by high erosive capacity. The central zone located in the Souss plain is dominated by medium erosive risk, and this indicates medium tendency to accumulate flood water (Table 2).

Topographic Wetness Index (TWI). The TWI is a physical index representing the effect of local topography on runoff and the direction of flow and its accumulation. The index indicates the accumulation of water in a watershed and thus applied in runoff modelling [29]. The index strongly represents areas of a watershed that are prone to flooding and was calculated as shown in [30]:

$$\mathrm{TWI} = \ln(\alpha/(\tan\beta)) \tag{2}$$

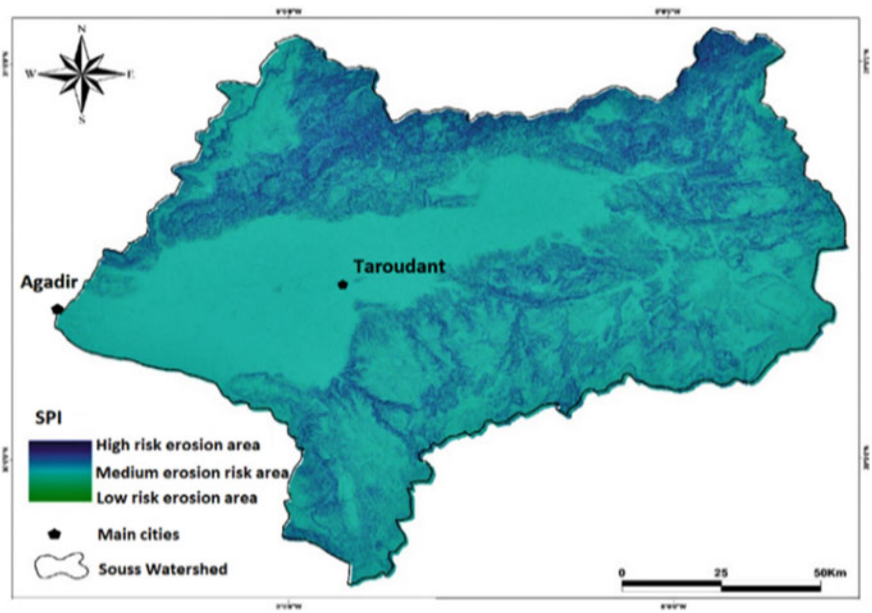


Fig. 7 SPI map in Souss watershed

Table 2 Erosion intensity distribution in the Souss watershed

Erosion intensity	Area (ha)	%
Low	417.908	23
Medium	474.496	27
High	904.580	50
Total	1796.984	100

Source ABHSM Monography report, 2014

where α is the cumulative drainage of the upstream area through a point (per unit contour length) and tan β is the slope angle at that point. The TWI map of the watershed was constructed with five intervals. The map shows that the areas where water is accumulated in the Souss watershed are illustrated by dark blue colour. We can see the central part, which is characterized by a low altitude and relatively flat areas, and four other small areas that present the dams in the Souss watershed (Fig. 8; (Table 3).

Normalized Difference Vegetation Index (NDVI). The Normalized Difference Vegetation Index is used to represent the relationship between vegetation and flooding in a basin [31]. In general, the NDVI values range from −1 to + 1, and the negative values represent surfaces other than vegetation cover, such as snow, water, or clouds. The positive values represent the vegetative cover of varying degrees. The higher the NDVI values, the denser is the vegetation. The reflectance of the bare soil is about

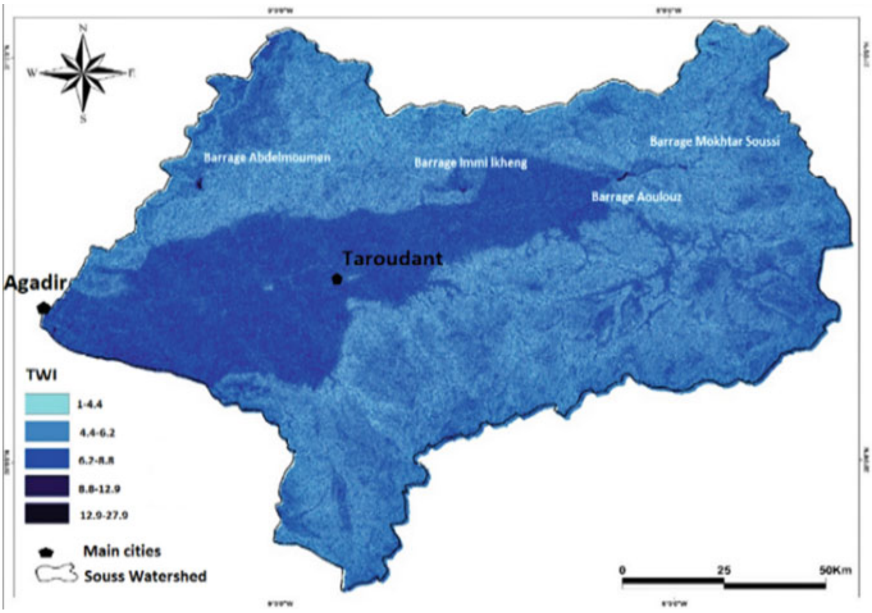


Fig. 8 TWI map in Souss watershed

Table 3 Main hydraulic dams of the Souss River

Dams	Entry into service	Capacity (million m^3)	Regulated volume (million m^3)
Barrage Abdelmoumen	1981	214	68.5
Barrage Aoulouz	1991	108	18
Barrage Immi lKheng	1993	11	5.5
Barrage Moukhtar Soussi	2001	50	45
Total	2001	383	137

Source ABHSM Monography report, 2014

the same in the red and near infrared, and the values are close to 0. The vegetation formations have positive NDVI values, generally between 0.1 and 0.7. The highest values correspond to the densest cover, and the NDVI map prepared with six classes (Fig. 9) was extracted from Landsat 8 OLI imagery. The NDVI values were calculated as [32]:

$$\mathrm{NDVI} = (\mathrm{PIR} - \mathrm{R})/(\mathrm{PIR} + \mathrm{R}) \tag{3}$$

The NDVI values shows that the basin is very poor in vegetation cover.

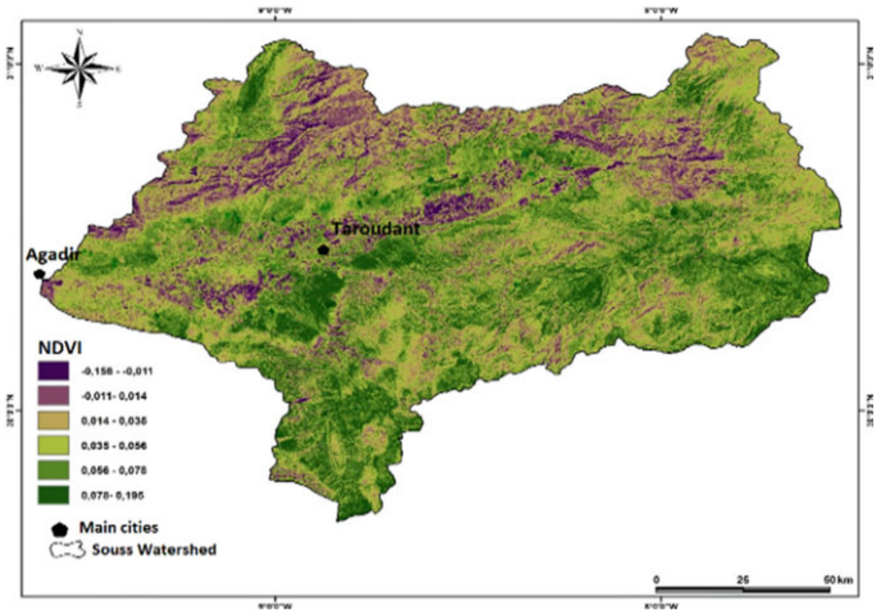


Fig. 9 NDVI map in Souss watershed

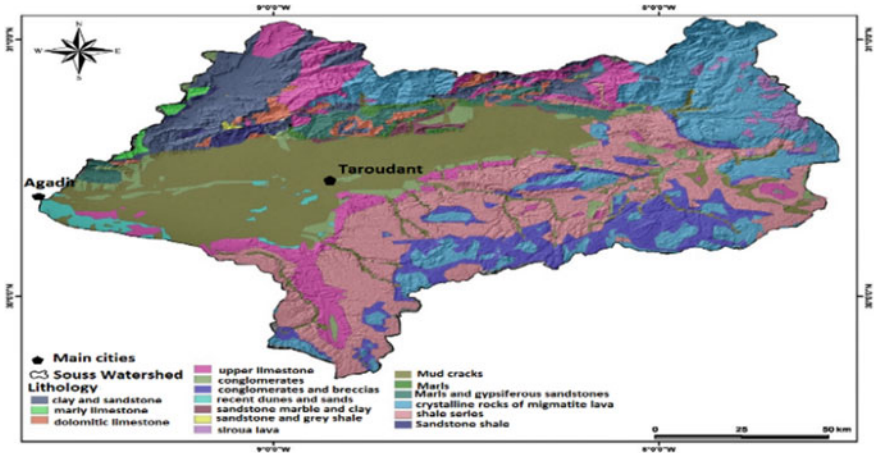


Fig. 10 Lithology map in Souss watershed

Lithology. The lithology is a major factor for understanding the variations of water flow and possible sedimentation of the watershed [33] in spatial and temporal dimensions. In addition, the petrographic formations, in terms of both erodibility and permeability, could also determine the flood hazard [34]. The lithological map is obtained by digitalizing a previous map [35] and converted into various lithological units. The predominance of alluvial silts and marls dominates in the plain of Souss (see Fig. 10).

Land use. The hydrological cycle, including interception, infiltration, and concentration of runoff behaviour, thus indirectly influences flooding in the watershed. The hydrological response and the magnitude of flood risk can be represented by the land use characteristics [36]. Land cover with forests and dense vegetation can infiltrate more water into the soil compared to other areas because higher the vegetation higher will be the interception of water and reduce the impact of water fall on the soil and thus reduces the runoff and increases the infiltration [25]. The land cover map was prepared using Landsat 8 OLI images and classified into four categories using supervised classification in ENVI 5.1 software. In the Souss watershed, the forest area is very limited and is located in the High Atlas Mountains. In the Souss plain and along the Souss River, there is a concentration of agricultural and economic activity and the development of transport infrastructures (Fig. 11).

Rainfall. Rainfall is the major factor determining the intensity of flooding [26]. Heavy rainfall within a shorter period of time (<6 h) usually results in flooding [26]. The probability of flood occurrence is proportional to the amount of rainfall [37]. In this study, we used 25 years of data (1993 to 2018) from three main stations for preparation of the precipitation maps. The precipitation map for the area was constructed using the IDW method with six class intervals. The torrential floods are characterized by a short response time, a high amplitude of flows, a large and varied solid load in nature with high levels of destructions (Figs. 12 and 13).

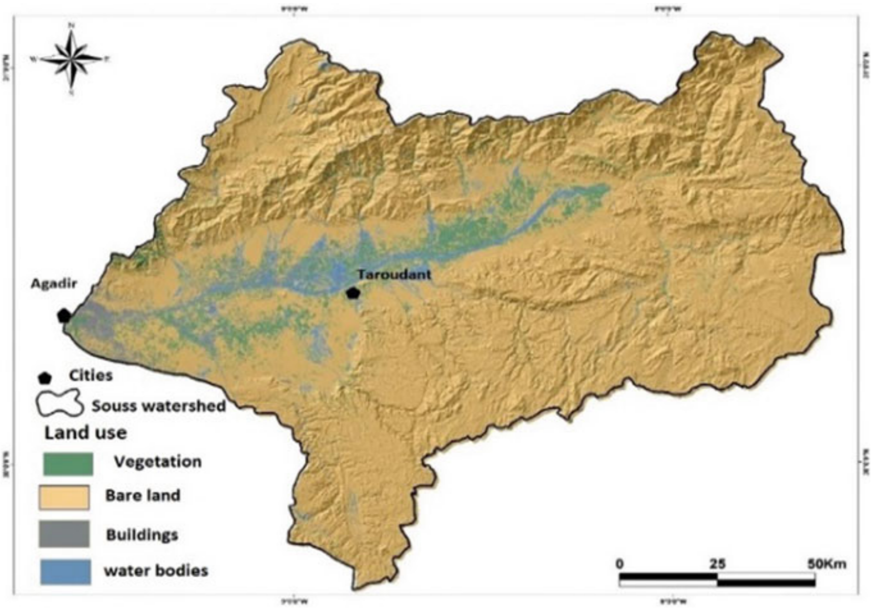


Fig. 11 Land use map in Souss watershed

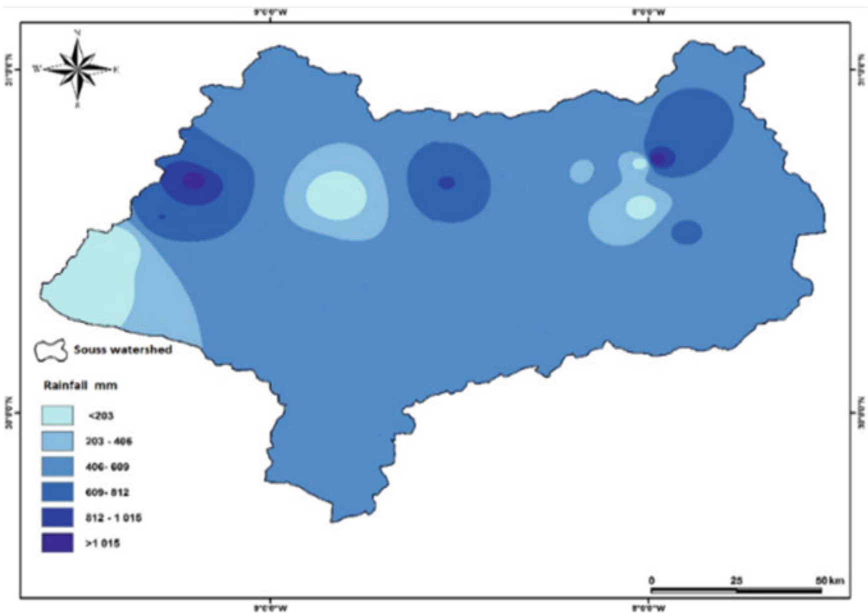


Fig. 12 Rainfall map during 25 years in the Souss watershed

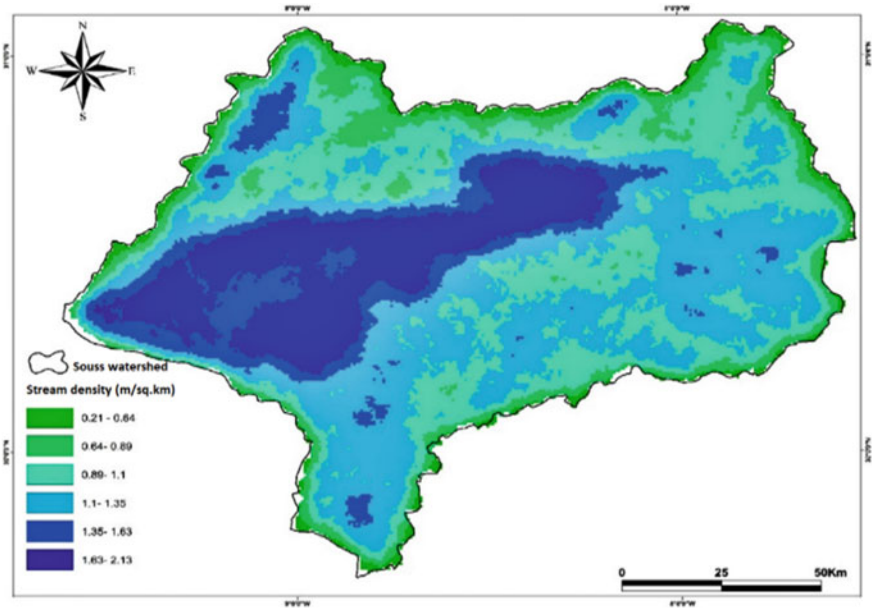


Fig. 13 Stream density map in the Souss watershed

Stream Density. Stream density refers to the ratio between the stream length (m) and the basin area (km^2) [38]. In general, flood-prone areas are mapped by taking into consideration the effect of each factor separately. However, the preparation of overall flood risk maps requires information about all the factors contributing to flooding. In that context, several factors can be combined by using the weighted overlay modelling method. Each of the raster layer representing the factors of flooding was assigned a weight in the analysis. The raster values are reclassified according to a common scale. The raster layers are overlaid by multiplying the value of each raster cell by its layer weight to arrive at the unique value for each cell. These values are assigned to new cells in an overall output layer. Assigning a weight to each raster in the overlay process controls the influence of different criteria in the model. Multiplying the weight of each layer by the suitability value of each cell gives a weighted suitability value. The weighted suitability values are summed for each overlay cell and then written to an output layer. The result is a flood risk index map that shows the area's most vulnerable to flooding in the selected watershed. The results of this analysis are presented in Figs. 14 and 15.

4 Results and Discussion

The results show that flat areas are subject to “high” and “very high” flood risks. The areas that are on higher elevation and in the upstream areas away from confluence and lower most points of the watershed are subject to “low” and “very low” flood risks. The most vulnerable areas are located along the central Souss watershed. These areas have a low slope angle and low altitude. As shown in the map, the plain contains several red areas that are closer to the main river of Souss. The road infrastructure is also strongly affected by these red zones as shown in Figs. 14, 15 and 16. The map shown in Fig. 14 presents the roads network and the network of important rivers in the Souss watershed. Where these two networks are very intertwined, the vulnerability to flooding increases especially in the urban areas. This increases the vulnerability of the urban network to flooding, which in turn has severe destructive effects such as the collapse of roads, the destruction of bridges, landslides, flooding of roads, massive deposits of silt (alluvial deposits, floating bodies), which can cause loss of human life.

A multi-criteria evaluation [39] is used by combining the factors with their weights, and then overlaying the thematic maps in a GIS environment, to create the map of flood-prone areas in the watershed; the results of this analysis step are presented in (Fig. 17). The map clearly shows that the areas at high risk of flooding

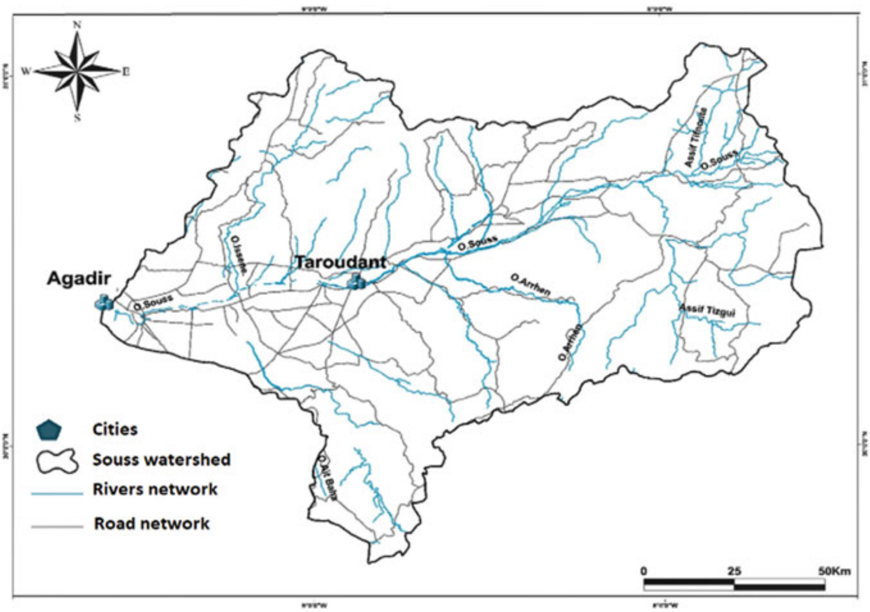


Fig. 14 Map showing the road network and main rivers in the Souss watershed

Fig. 15 Road over the Oughri river in Ouled Berhil village, 40 km from Taroudant has lost a huge part due to floods (*Source* ABHSM, May 2016)

Fig. 16 Damaged road due to floods, Taroudant (*Source* ABHSM, 2019)

are concentrated in the central part located in the plain of watershed with low altitude, a low degree of slope, a great proximity to rivers including the main river of Souss and which run through the road and urban infrastructure.

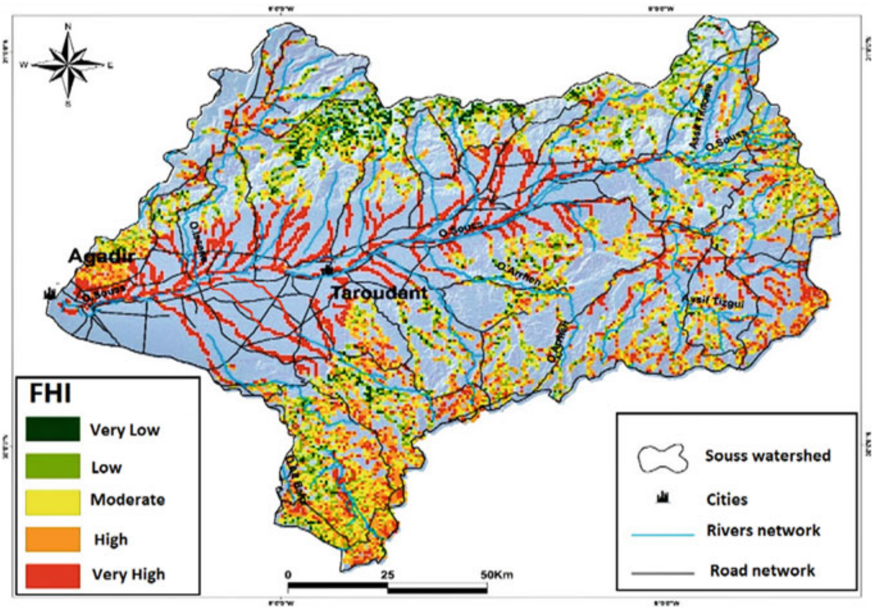


Fig. 17 Map showing flood areas in Souss watershed

The results of the flood risk maps were validated using the historical flood events in selected locations which were collected based on field information from the Souss Massa Draa Hydraulic Basin Agency. These historical floods were superimposed on the output modelled map. It was noticed that all historical flood points are located on the high and very high vulnerability zones (Fig. 18). The results clearly showed that the 10 factors used for determination of flood vulnerability zones using the model clearly correspond to the historical flooding events in the region.

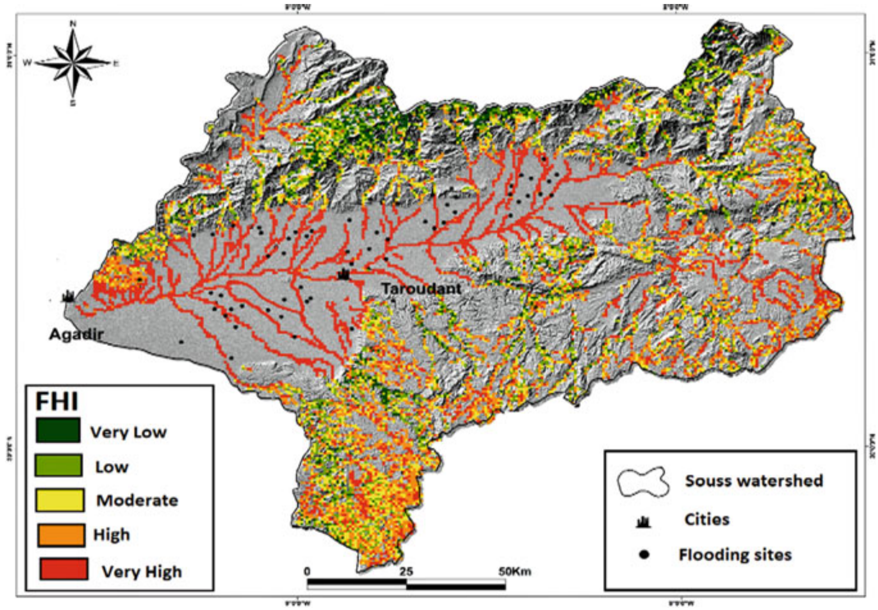


Fig. 18 Map showing the historical flood sites and the flood zones delimited by our model in the Souss watershed

5 Conclusion

The flood risk maps are very important for flood prevention and risk management. The flood risk maps help the decision-makers, water resource managers and planners to assess the potential risks and accordingly plan and implement flood protection and possibly flood prevention measures. The multi-criteria assessment demonstrated the applicability of the thematic layers combined to produce unique flood risk maps for decision-making for risk management and prevention. The analysis and subsequent validation clearly showed the applicability and suitability of 10 thematic layers that represent most important factors responsible for flood risk. However, longer-term historical flood information in the Souss watershed according to the ABHSMD could be used in the future analysis to further advance the validation of the model.

The results of this study clearly showed the validity of the methodology for mapping of flood risk. The study also highlighted the most prominent factors responsible for flood occurrence. It is evidenced from the analysis that the vulnerability of flood occurrence increases with high-intensity rainfall, closeness of the area to the river, abstractions of water flow by urban infrastructure, low-lying plains with poor vegetation. The methodology and approach described above hold promise to develop spatial maps for comprehensive and integrated planning of risk management measures at the basin scale.

In addition, the information obtained from this study could be used in future research and to further investigate the effects of various factors responsible for flooding and to develop new models for flood risk assessment. The study clearly shows that the floods in Souss watershed are heavily influenced by the intensity and irregularity of rainfall regime in Souss watershed [7], which is considered as the most determining factor for occurrence of floods, followed by the absence of vegetation cover. The irregularity of rainfall has been accentuated these last decades due climate change. In fact, several national and international studies [40–42] confirmed that Morocco is one of the most vulnerable regions vulnerable to climate change. The rainfall variability of the Souss watershed favours the occurrence of floods, due to of heavy rainfall in a short time and lack of the vegetation cover in the area.

Such improved flood risk maps with combination of information could be used to implement proactive risk management measures than spending signification resources to reactive emergency response to flooding. The proactive flood risk management is several times economical than reactive emergency response measures. Thus, the analysis and mapping could be an effective tool contributing to the proactive flood risk management for the Souss watershed and also many other similar areas or watershed.

However, flood risk management has to adopt a comprehensive approach involving multiple sectors and prepare integrated plans that could protect all economic sectors such as agriculture, fisheries and tourism that are prevalent in this region. “The path is long, the roots are bitter, but the fruit is sweet” [43]. Such is summarized the laborious process of study of the area of the “Souss watershed”, which started with the exploration of the area in question, through the collection, processing, analysis and exploitation of data, to finally give rise to the demonstration of these important results. This study focusing on the Souss watershed could be an effective reference for other regions and watersheds in Morocco and the method can be applied in semi-arid and arid climate in several regions in the world, by adding their own characteristics and information.

Though the methodology is robust enough to identify flood-prone areas and plan flood risk management measures, the methodology could be further advanced to improve the accuracy in identification of flood-prone areas. The future work may look in to identification of additional factors responsible for flooding which is obviously different for various regions. Inclusion of additional layers on detailed soil types and its properties, and population distribution and other demographic and land use features could contribute to improve the accuracy in identification of flood-prone areas. Further analysis should identify and categorize factors of flood hazard risks and factors contributing to vulnerability. These factors may include land use, population distribution, infrastructures, such as roads and bridges in the watershed. Including multiple factors in the analysis could increase the accuracy and relevance of results to implement flood mitigation measures.

Overall, the flood risk mapping approach presented in this paper can be used to plan and accordingly to avoid significant damage to road and other infrastructure facilities in urban areas and loss of assets and livelihoods. To achieve flood preparedness and prevention, the flood risk maps could be integrated with early warning systems that is able to provide advance information about the timing and intensity of flooding. The advance early warning system with sufficient lead time could affectively protect the most vulnerable populations from loss of their assets and livelihoods.

Acknowledgements and Declaration The lead author Jada El kasri has conceptualized the technical work, developed and validated the approach and methodology. Abdelaziz Lahmili, Ahmed Bouajaj and Halima Soussi provided suggestions to the paper and reviewed the paper and provided comments. The manuscript is read and approved by all authors. The authors acknowledge the Souss Massa Hydraulic Basin Agency in Agadir for making available climate data and reports. There is no conflict of interest for the authors.

References

1. Kim D, Jung HS, Baek W (2016) Comparative analysis among radar image filters for flood mapping. J Korean Soc Surv Geodesy Photogrammetry and Cartography 34:43–52
2. Novelo-Casanova DA, Rodrıguez-Vangort F (2016) Flood risk assessment. Case of study: Motozintla de Mendoza, Chiapas, Mexico. Geomat Nat Haz Risk. 7:1538–1556
3. Shen G, Hwang SN (2019) Spatial—temporal snapshots of global natural disaster impacts Revealed from EM-DAT for 1900–2015. Geomat Nat Hazards Risk 10:912–934
4. Klaus S, Kreibich H, Merz B, Kuhlmann B, Schroter K (2016) Large-scale, seasonal flood risk analysis for agricultural crops in Germany. Environ Earth Sci 75:1289
5. Chapi K, Singh VP, Shirzadi A, Shahabi H, Bui DT, Pham BT, Khosravi K (2017) A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ Model Softw 95:229–245
6. Roy P, Pal SC, Chakrabortty R, Chowdhuri I, Malik S, Das B (2020) Threats of climate and land use change on future flood susceptibility. J Clean Prod 272:122757
7. El Kasri J, Lahmili A, Soussi H, Jaouda I, Bentaher M (2021) Trend analysis of meteorological variables: rainfall and temperature. Civil Eng J 7(11):1868–1879
8. Chang HS, Chen TL (2016) Spatial heterogeneity of local flood vulnerability indicators within flood-prone areas in Taiwan. Environ Earth Sci 75:1484
9. Meyer M et al. (2001) Satellite remote sensing techniques used in archaeological research in Luristan, Western Iran. In: Proceedings of the 1st EARSeL workshop on remote sensing for developing Countries. Gent, Belgium, pp 297–303
10. Grimaldi S, Petroselli A, Arcangeletti E, Nardi F (2013) Flood mapping in ungauged basins using fully continuous hydrologic–hydraulic modeling. J Hydrol 487:39–47
11. Papaioannou G, Loukas A, Georgiadis C (2013) The effect of riverine terrain spatial resolution on flood modeling and mapping. In: Proceedings of the first international conference on remote sensing and geoinformation of environment. Bellingham, International Society for Optics and Photonics
12. Chini M, Giustarini L, Matgen P, Hostache R, Pappenberger F, Bally P (2014) Flood hazard mapping combining high resolution multi-temporal SAR data and coarse resolution global hydrodynamic modelling. In: Proceedings of the IEEE geoscience and remote sensing symposium. New York (NY), IEEE
13. Curebal I, Efe R, Ozdemir H, Soykan A, S€onmez S (2016) GIS-based approach for flood analysis: case study of Kecidere ¸ flash flood event (Turkey). Geocarto Int 31:355–366
14. Disyacitta A, Nurul H, Zahrotul M, Dwi N (2017) Spatial analysis for potential water catchment areas using GIS: weighted overlay technique. In: IOP conference series: earth and environmental science. vol 98. pp 012054. https://doi.org/10.1088/1755-1315/98/1/012054

15. Basharat M, Shah H, Hameed N (2016) Landslide susceptibility mapping using GIS and weighted overlay method: a case study from NW Himalayas, Pakistan. Arabian J Geosci 9. https://doi.org/10.1007/s12517-016-2308-y
16. Pelin OS, Tarhan C (2016) Detection of flood hazard in urban areas using GIS: Izmir case. Proc Technol 22:373–381. ISSN 2212–0173
17. Karimi H, Bengin MA, Herki B, Gharibi S, Tehrani SH, Kakhani A (2020) Identifying public parking sites using integrating GIS and ordered weighted averaging approach in Sanandaj city, Iran. J Critical Rev 7(4). ISSN-2394–5125
18. Das S, Gupta A (2021) Multi-criteria decision based geospatial mapping of flood susceptibility and temporal hydro-geomorphic changes in the Subarnarekha river, India. Geosci Front
19. Saaty TL (1990) how to make a decision: the analytic hierarchy process. Eur J Oper Res 48(1):9–26
20. Bui DT, Tuan TA, Klempe H, Pradhan B, Revhaug I (2016) b. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13(2):361–378
21. Sanyal J, Lu XX (2004) Application of remote sensing in flood management with special reference to Monsoon Asia: a review. Nat Hazards 33:283–301
22. Khosravi K, Nohani E, Maroufinia E, Pourghasemi HR (2016) A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique. Nat Hazards 83:947–987
23. Pradhan B (2010) Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. J Spat Hydrol 9(2)
24. Khosravi K, Pourghasemi HR, Chapi K, Bahri M (2016) Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical index, and weighting factor models. Environ Monit Assess 188:656
25. Tehrany MS, Pradhan B, Jebur MN (2014) Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. J Hydrol 512:332–343
26. Cao C, Xu P, Wang Y, Chen J, Zheng L, Niu C (2016) Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas. Sustainability 8(9):948
27. Poudyal CP, Chang C, Oh HJ, Lee S (2010) Landslide susceptibility maps comparing frequency ratio and artificial neural networks: a case study from the Nepal Himalaya. Environ Earth Sci 61(5):1049–1064
28. Turoglu H, Dolke _ I (2011) Floods and their likely impacts on ecological environment in the Bolaman river basin (ORDU, TURKEY). Res J Agric Sci 43(4):167–173
29. Beven KJ (2011). In: Rainfall-runoff modelling: the primer. Wiley
30. Beven K, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol Sci J 24(1):43–69
31. Tehrany MS, Pradhan B, Jebur MN (2013) Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. J Hydrol 504:69e79
32. Tucker CJ, Justice CO, Prince SD (1986) Monitoring the grasslands of the Sahel 1984–1985. Int J Remote Sens 7:1571–1581
33. Miller DD, McPherson JG, Covington TE (1990) Fluviodeltaic reservoir. South Belridge Field
34. Stefanidis S, Stathis D (2013) Assessment of flood hazard based on natural and anthropogenic factors using analytic hierarchy process (AHP). Nat hazards 68(2):569–585
35. Malki M, Choukr-Allah R, Bouchaou L, Hirich A, Brahim YA, Krimissa S, Hssaisoune M (2016) Assessment of groundwater quality: impact of natural and anthropogenic contamination in Souss-Massa River Basin. In: The handbook of environmental chemistry book series, HEC, vol 53

36. Rahmati O, Samani AN, Mahdavi M, Pourghasemi HR, Zeinivand H (2015) Groundwater potential mapping at Kurdistan region of Iran using analytic hierarchy process and GIS. Arabian J Geosci 8(9)
37. Todini F, De Filippis T, De Chiara G, Maracchi G, Martina M, Todini E (2004) Using a GIS approach to asses flood hazard at national scale. In: Proceedings of the European geosciences union, 1st General Assembly, April, pp 25e30
38. Andrew E, Jason J, Steven G, Matthew F (2013) Potential stream density in mid-Atlantic U.S. Watersheds. PloS one 8:e74819. https://doi.org/10.1371/journal.pone.0074819
39. Goepel KD (2018) Implementation of an online software tool for the analytic hierarchy process (AHP-OS). Int J Anal Hierarchy Process 10(3):469–487
40. Khattabi A, Chriyaa A, Hammani A, Brahim M (2014) Vulnérabilités climatiques et stratégies de dévelopment: synthese et recommandations stratégiques pour une prise en compte du risque « climat » dans les politiqueset stratégies sectorielles. https://doi.org/10.13140/RG.2.1.3081.2562
41. Change C (2007) Synthesis report: contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change core writing team. In: IPCC, Geneva, Switzerland, 104. Geneva, Switzerland, IPCC, pp 104
42. Bouhali H, Payen J (2019) Etude d'impact Des Changements Climatiques Sur Les Ressources En Eau et Les Risques d'inondations Dans La Vallée d'Arghen -Bassin de Souss-Massa, Experts-Solidaires Septe. Ecole Hassanya Des Travaux Publics Avec Jean Payen. Experts-Solidaires 116. Available online. https://experts-solidaires.org/wp-content/uploads/2020/01/Rapport-Changement-Climatique-et-Ressources-en-Eau-dans-la-vallée-dArghen
43. "The path is long, the roots are bitter, but the fruit is sweet "Master Pham Xuân Tong (founder of Qwan Ki Do)

Computational Analysis of a Mobile Path-Planning via Quarter-Sweep Two-Parameter Over-Relaxation

A'Qilah Ahmad Dahalan and Azali Saudi

Abstract Over the years, self-reliant navigation has risen to the forefront of research topics. Improving the path-planning competencies is an extremely important component in achieving excellent autonomous navigation. This paper describes a refinement of the proficiency of mobile path-planning through a computational approach, i.e. the quarter-sweep two-parameter over-relaxation (QSTOR), to solving path-planning problems iteratively. The solution of Laplace's equation (otherwise known as the harmonic functions) is the source for producing the potential function of the configuration space of the mobile robot. Numerical experiments illustrate that, in a given environment, a mobile robot is able to steer towards a particular destination with a smooth and ideal path from any beginning location. Furthermore, it is shown that in terms of the iterations number and computational time, the QSTOR iterative technique outperforms its predecessors in addressing mobile path-planning issues.

Keywords Laplace's equation · Finite difference method · Accelerated over-relaxation · Path navigation · Optimal route · Obstacle avoidance · Quarter-sweep iterative techniques

1 Introduction

The robotics discipline is gaining traction in our daily lives as well as in various domains of modern industrial and cyber-physical automation. With the ability to embed intelligence into robots becoming more widely available, identifying the

A. A. Dahalan (✉)
CONFIRM Centre for SMART Manufacturing, University of Limerick, Limerick, Ireland
e-mail: A.Qilah@ul.ie

Centre for Defence Foundation Studies, Universiti Pertahanan Nasional Malaysia, Kuala Lumpur, Malaysia

A. Saudi
Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_23

optimal solutions in the execution of any task, such as for path-planning and navigation would be easily accomplished. These kinds of tasks could be said as one of the most complex challenges in intelligence robots. In the direction of constructing an autonomous mobile robot, it is important for the robot to be competent and accurate in creating a route as well as be collision-free. Practical algorithms concerning this difficulty have great exploitation such as in computer animation [1], robotics manufacturing [2], architectural design [3], including security, defence, and surveillance [4, 5].

The aim of this paper is to use numerical potential functions on simulating a driving point-robot in the configuration space analogously by heat distribution [6]. The employment of such a heat transfer paradigm results in an environment with no local minima, which give hugely beneficial for robot path-planning. Laplace's equation is utilized to depict the analogy of heat distribution across the experiments. The "temperature values" for the path creation model in the environment, referred to as configuration space (C-space) are characterized by the solution of Laplace's equation, i.e. the harmonic functions. To solve these functions, a variety of approaches have been explored, while numerical techniques are most typically used due to their fast-processing mechanisms and proficiency in solving the problem. This paper conducted a number of tests to examine the performance of proposed accelerated algorithms in generating mobile robot paths. The objective behind this study is to examine and verify the efficacy of the proposed algorithm before attempting to integrate the algorithm into a real robot for a subsequent study.

2 Path-Planning Structure

Path-planning, in general, allows an autonomous vehicle or a robot to discover the shortest and safest most obstacle-free path from a starting point to a destination. Indoor mobile robot path navigation can be achieved in many different ways. A path navigation algorithm for an identified environment can certainly yield a series of nodes for a robot to trail. Typically, a grid of a predetermined size is created to evaluate different algorithms, showing where "passable" is on the C-space. It is reasonable to assume that the robot can traverse all of the grid's boundaries.

The structure of this experiment is based on the use of a point-robot to simulate the motion within the recognized C-space. The robot's route is determined using a heat transfer analogy in which the target point (with the lowest potential value) serves as a heat-pulling sink. While every wall and obstacle (with the highest potential value) is regarded as a heat source that should always be set as constant. In compliance with the heat transfer behaviour, the heat will flow from a higher-temperature region towards a lower-temperature region, completing the C-space. This event is represented by harmonic function values, which will result in so-called heat flux lines flowing/streaming towards the region with the lowest potential value, i.e. the sink. The path line for the robot to traverse across the C-space was built out in this arrangement, by following the heat flux line produced. The implementation of the harmonic function prevents the event of local minima and can guide the robot to avoid obstacles in the environment [7].

2.1 Harmonic Functions

A Laplace's equation-satisfying function is known as a harmonic function provided in the domain $\Omega \subset R^n$. The borderline of every wall, each obstacle in the region, primary points, and target points are all contained within the boundary of Ω for the development of the robot path. Consider Laplace's equation below with x_i is the ith coordinates in the Cartesian plane, and n is the dimension.

$$\nabla^2\phi = \sum_{i=1}^{n} \frac{\partial^2\phi}{\partial x_i^2} = 0. \tag{1}$$

By using the numerical approach, i.e. Jacobi or Gauss–Seidel (GS), Laplace's Eq. (1) could be adequately solved. The harmonic function has been shown that it abides by the min–max principle, which implies it prevents the formation of spurious local minima excluding the target point and typically creates a smooth path [8]. For this reason, the harmonic potential technique is a viable and appealing decision for robot path-planning. Most often, conventional methods [9–11] are used to solve the Laplace equation. Equation (1) in this paper was solved using the quarter-sweep iterative approach to improve the acceleration of the computational execution.

A global approach is used to measure the harmonic potentials of the robot C-space for path-planning problems. The trail lines for a robot to move along from start to end location without encountering any obstacles are mapped using potential solutions for Eq. (1). As mentioned earlier, obstacles and walls are viewed as current sources whiles the target point is to be the sink. The Dirichlet boundary conditions provide boundary values. Following that, by performing a standard gradient descent search (GDS) on the potential field, a sequence of potential points with lower values is found, progressing to the point with the lowest potential value, which is the target location.

Altogether, this paper attempts to replicate the stated path-planning paradigm, defining the solution of Laplace's equation over the resemblance of temperature (for the potential) and heat flow (for the path line). The experimentation takes place on a two-dimensional domain with assorted shapes of obstacles, along with the walls. To address Eq. (1) in gaining potential values for each node, the quarter-sweep two-parameter over-relaxation (QSTOR) scheme is employed. The existing technique (i.e. families of over-relaxation methods) was also measured for comparison to analyze the competence of the proposed scheme.

3 Materials and Techniques

From Eq. (1), the two-dimensional Laplace's equation is given as

$$\nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0. \tag{2}$$

The Laplacian operator is implied by ∇^2. To compute Eq. (2) using a numerical method, it should be discretized over the simplest five-point finite difference approximation (5P-FDA). For two-dimensional Laplace's Eq. (2), let $U_{i,j}$ approaches the solution of u along the grid point $\left(x_i, y_j\right)$, hence the discretization of these Laplace equations by conventional five-point stencil is written as

$$U_{i-1,j} + U_{i+1,j} + U_{i,j-1} + U_{i,j+1} - 4U_{i,j} = 0. \tag{3}$$

The iterative routine for Laplace's Eq. (2) is implying swapping the node value continuously with the median of its four neighbours. In parallel, all nodes in the grid point will be computed using Eq. (3), this action is called full-sweep (FS) iteration (see Fig. 1a). Abdullah [12] later initiated the explicit decoupled group, which was then known as the half-sweep (HS) approach. This method demonstrates an effective technique for solving PDEs [13–16]. Since the HS technique yielded such promising results, Othman and Abdullah [17] came out with an improved approach, namely modified explicit group, also known as quarter-sweep (QS). Figure 1 indicates the computational mesh of each sweep technique, where only black points are evaluated for the whole iteration cycle. In the mesh region, only half and a quarter of the node points are calculated using HS and QS schemes, respectively. Rationally, this signifies the reduction of computational time on each iteration. Figure 2 shows the computational stencils of each technique. It is observed that the HS iteration is primarily based on rotated 5P-FDA in solving the Laplace equation, given as

$$U_{i-1,j-1} + U_{i+1,j-1} + U_{i-1,j+1} + U_{i+1,j+1} - 4U_{i,j} = 0. \tag{4}$$

3.1 *Conceptualization of the QS Method*

The implementation of the QS iterative scheme will compute only one out of four of the nodal points at one time (see Fig. 1c) during the iteration process in the C-space. Consequently, it will decrease the computational complexity drastically, i.e. roughly 75%. The QS approximation equation precisely skipped two nodal points from the mesh space (see Fig. 2c). Therefore, the formula of QS five-point approximation can be written as

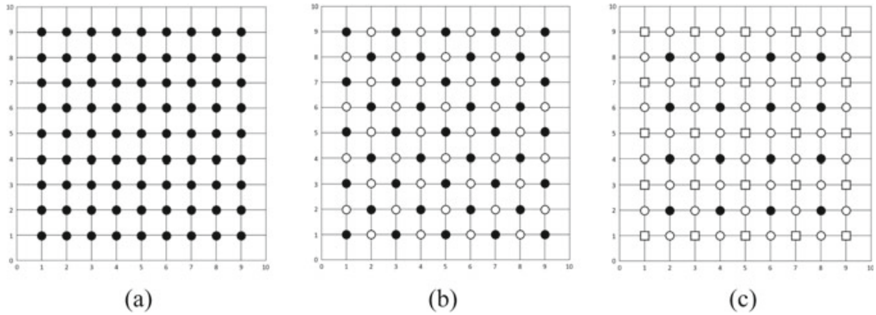


Fig. 1 Computational mesh of **a** FS, **b** HS, and **c** QS technique

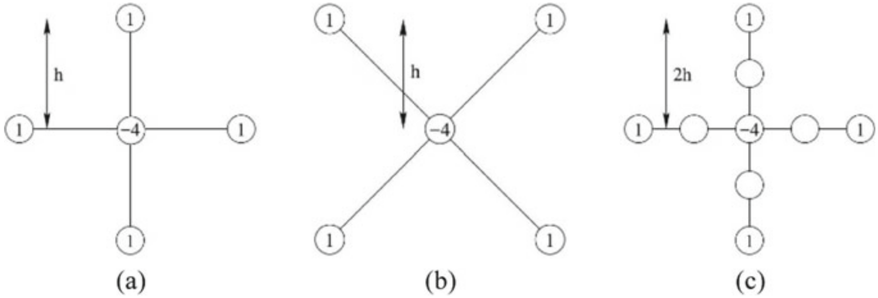


Fig. 2 Computational stencil of **a** FS, **b** HS, and **c** QS technique

$$U_{i-2,j} + U_{i+2,j} + U_{i,j-2} + U_{i,j+2} - 4U_{i,j} = 0. \tag{5a}$$

Considering finite difference from Eq. (5a), the GS iterative technique for QS can be rewritten and denoted as

$$U_{i,j}^{(k+1)} = \frac{1}{4}\left[U_{i-2,j}^{(k+1)} + U_{i+2,j}^{(k)} + U_{i,j-2}^{(k+1)} + U_{i,j+2}^{(k)}\right]. \tag{5b}$$

Successive over-relaxation (SOR) is basically a variant of the GS technique. When implanted SOR approach into Eq. (5b) by appending a weighted parameter ω [18], the QSSOR iterative scheme is given as

$$U_{i,j}^{(k+1)} = \frac{\omega}{4}\left[U_{i-2,j}^{(k+1)} + U_{i+2,j}^{(k)} + U_{i,j-2}^{(k+1)} + U_{i,j+2}^{(k)}\right] + (1-\omega)U_{i,j}^{(k)}. \tag{6}$$

Be noted that whenever $\omega = 1$, then the SOR approach is in fact simplified to the GS method.

The accelerated over-relaxation (AOR) fundamentally is a simplification of the SOR technique with additional optimal parameters, denoted as ω and ω' in this paper. To execute the AOR scheme as proposed in [19], the node points of $u_{i-1,j-1}^{(k+1)}$ and

$u_{i+1,j-1}^{(k+1)}$ are interchanged to $u_{i-1,j-1}^{(k)}$ and $u_{i+1,j-1}^{(k)}$ respectively, as well as inserting the $\frac{\omega'\left(u_{i-1,j-1}^{(k+1)}-u_{i-1,j-1}^{(k)}\right)}{4}$ and $\frac{\omega'\left(u_{i+1,j-1}^{(k+1)}-u_{i+1,j-1}^{(k)}\right)}{4}$ nodes into Eq. (6). Now, the new scheme of QSAOR is provided as

$$\begin{aligned} U_{i,j}^{(k+1)} &= \frac{\omega'}{4}\left[U_{i-2,j}^{(k+1)} - U_{i-2,j}^{(k)} + U_{i,j-2}^{(k+1)} - U_{i,j-2}^{(k)}\right] \\ &+ \frac{\omega}{4}\left[U_{i-2,j}^{(k)} + U_{i+2,j}^{(k)} + U_{i,j-2}^{(k)} + U_{i,j+2}^{(k)}\right] + (1-\omega)U_{i,j}^{(k)}. \end{aligned} \tag{7}$$

Meanwhile, the two-parameter over-relaxation (TOR) technique is indeed a deduction from the AOR scheme. The main intention of this technique is to improve the convergence speed, ergo of it consists three different relaxation parameters: ω, ω', and ω''. Thus, the QSTOR iterative scheme is

$$\begin{aligned} U_{i,j}^{(k+1)} &= \frac{\omega'}{4}U_{i,j-2}^{(k+1)} + \frac{\omega''}{4}U_{i-2,j}^{(k+1)} + \frac{\omega}{4}\left(U_{i,j+2}^{(k)} + U_{i+2,j}^{(k)}\right) \\ &+ \left(\frac{\omega-\omega'}{4}\right)U_{i,j-2}^{(k)} + \left(\frac{\omega-\omega''}{4}\right)U_{i-2,j}^{(k)} + (1-\omega)U_{i,j}^{(k)}. \end{aligned} \tag{8}$$

The uncertainty of relaxation parameter values has resulted in the minimum iteration counts. Previous researchers [19, 20] specified that the values of ω' and ω'' are generally chosen remain near to the SOR ω value. The computation is then recurrent for a range of $1 \le \omega < 2$. So as to discover the optimum value, the relaxation parameter values are individual for each sweep case, as certain values are not converged in some cases (not every value is converged in every case). Additionally, as the values of each parameter are predetermined before execution, the impact of complexity on determining the value of parameters on the entire computation is unaffected. It will certainly shift if the few ranges of parameter values are set in the computation algorithm. The implementation of the QSTOR scheme to solve Laplace's problem (2) is described in Algorithm 1.

Algorithm 1. QSTOR iterative scheme	
i.	Set up the C-space through the designated start and target points
ii.	Initialising starting point U, $\varepsilon \leftarrow 10^{-15}$, $iteration \leftarrow 0$
iii.	For every • node points, calculate $U_{i,j}^{(k+1)} \leftarrow \frac{\omega'}{4}U_{i,j-2}^{(k+1)} + \frac{\omega''}{4}U_{i-2,j}^{(k+1)} + \frac{\omega}{4}\left(U_{i,j+2}^{(k)} + U_{i+2,j}^{(k)}\right) + \left(\frac{\omega-\omega'}{4}\right)U_{i,j-2}^{(k)} + \left(\frac{\omega-\omega''}{4}\right)U_{i-2,j}^{(k)} + (1-\omega)U_{i,j}^{(k)}.$

(continued)

(continued)

Algorithm 1. QSTOR iterative scheme	
iv.	Compute the remaining □ node points via the direct method $U_{i,j}^{(k+1)} \leftarrow \frac{1}{4}\left[U_{i-1,j-1}^{(k+1)} + U_{i+1,j-1}^{(k+1)} + U_{i-1,j+1}^{(k)} + U_{i+1,j+1}^{(k)}\right]$, and ∘ node points by using $U_{i,j}^{(k+1)} \leftarrow \frac{1}{4}\left[U_{i-1,j}^{(k+1)} + U_{i+1,j}^{(k)} + U_{i,j-1}^{(k+1)} + U_{i,j+1}^{(k)}\right]$
v.	Verify the convergence test for $\varepsilon \leftarrow 10^{-15}$, then perform GDS to create a path towards the target. Otherwise, go back to step (iii)

4 Experiments and Results

There are four different C-spaces (with assorted obstacles) over four separate mesh sizes through the simulation experiments in this study. Although no specific potential values were appointed to any starting position, the target point was placed at the lowermost temperature values. During the initial setting, every obstacle and wall were assigned with the highest potential value where boundary values are described by the Dirichlet boundary conditions. The free spaces in the environment were made to be zero potential.

The computational process was carried out using an AMD A10-7400P Radeon R6 with 10 Compute Cores 4C + 6G running at 2.50 GHz and 8 GB of RAM. Provided that the state for stopping criteria is satisfied, the process of iteratively measuring potential values at each point continues. The iteration loop will be terminated, where the variance of the computational values was extremely small, (i.e. 1.0^{-15}), if the potential values do not show any further changes. This level of precision was necessary for the solutions to avoid saddle points, which are flat areas that fail to produce routes.

The iteration number and the execution time for every computational approach are, respectively, given in Tables 1 and 2. As compared to other suggested techniques, the QSTOR iterative scheme has been proven that it is significantly faster. It is demonstrated that, in terms of iteration number, the QSTOR outperformed the QSAOR (approximately by 5–12%) and QSSOR (approximately by 15–28%). On the other hand, the QSTOR decreases QSSOR from 10 to 18% and QSAOR from 9 to 20% in terms of execution time.

4.1 Discussion

The moment the potential values were gained, the route was constructed by carrying out the steepest descent search following the initial points to the specified destination.

Table 1 Findings of the proposed schemes for iteration number

	Methods	$N \times N$			
		300	600	900	1200
Condition 1	FSSOR	1728	8117	17,831	31,346
	FSAOR	1591	7529	16,594	28,984
	FSTOR	1656	7815	17,199	27,895
	HSSOR	837	4108	9086	15,892
	HSAOR	759	3803	8420	14,768
	HSTOR	797	3949	8721	14,234
	QSSOR	351	2078	4632	8113
	QSAOR	348	1913	4280	7508
	QSTOR	*344*	*1992*	*4448*	*7279*
Condition 2	FSSOR	2228	8776	19,254	33,558
	FSAOR	2006	7973	17,538	30,573
	FSTOR	1893	7553	16,642	29,008
	HSSOR	1071	4438	9813	17,149
	HSAOR	944	4023	8924	15,614
	HSTOR	877	3811	8461	14,813
	QSSOR	452	2229	5014	8771
	QSAOR	430	2007	4542	7976
	QSTOR	*414*	*1890*	*4305*	*7558*
Condition 3	FSSOR	3624	14,644	33,004	57,484
	FSAOR	3236	13,165	29,680	51,738
	FSTOR	2843	11,685	26,393	46,021
	HSSOR	1780	7445	16,856	29,418
	HSAOR	1568	6681	15,149	26,456
	HSTOR	1349	5909	13,463	23,523
	QSSOR	828	3769	8624	15,061
	QSAOR	698	3366	7740	13,545
	QSTOR	*512*	*2960*	*6856*	*12,023*
Condition 4	FSSOR	2507	9868	21,654	37,762
	FSAOR	2288	9025	19,840	34,601
	FSTOR	2067	8217	18,052	31,519
	HSSOR	1212	5000	11,036	19,288
	HSAOR	1097	4555	10,098	17,670
	HSTOR	967	4141	9180	16,085
	QSSOR	555	2502	5638	9873
	QSAOR	467	2287	5148	9030
	QSTOR	*427*	*2066*	*4676*	*8215*

Table 2 Findings of the proposed schemes for execution time (in second)

	Methods	N × N			
		300	600	900	1200
Condition 1	FSSOR	8.13	227.95	1134.25	3728.92
	FSAOR	8.61	230.17	1148.87	3692.74
	FSTOR	7.60	233.91	1188.08	3565.09
	HSSOR	2.39	81.24	404.15	1375.27
	HSAOR	1.72	73.76	369.91	1247.65
	HSTOR	2.55	84.84	413.84	1335.52
	QSSOR	0.39	14.99	81.55	293.92
	QSAOR	0.56	15.83	84.47	292.46
	QSTOR	*0.38*	*16.46*	*87.40*	*279.95*
Condition 2	FSSOR	10.69	251.72	1270.23	4077.22
	FSAOR	10.27	248.24	1226.66	3976.33
	FSTOR	9.39	233.83	1194.50	3732.02
	HSSOR	2.95	86.77	445.70	1423.27
	HSAOR	2.75	76.79	403.25	1263.63
	HSTOR	2.70	82.42	401.42	1326.65
	QSSOR	0.64	16.69	90.03	313.44
	QSAOR	0.56	16.68	89.98	314.14
	QSTOR	*0.52*	*15.19*	*85.08*	*287.87*
Condition 3	FSSOR	16.22	427.27	2190.45	7432.68
	FSAOR	18.66	418.45	2073.25	7254.02
	FSTOR	15.20	369.55	1927.30	6300.13
	HSSOR	5.16	154.79	783.72	2634.52
	HSAOR	4.80	137.18	721.94	2300.84
	HSTOR	4.30	135.81	661.90	2262.25
	QSSOR	0.92	30.04	166.12	567.28
	QSAOR	1.08	29.24	161.76	570.33
	QSTOR	*0.77*	*25.35*	*144.71*	*488.66*
Condition 4	FSSOR	11.02	281.85	1441.47	4853.57
	FSAOR	12.52	281.78	1423.54	4743.21
	FSTOR	10.91	255.82	1292.23	4269.42
	HSSOR	3.58	102.16	510.22	1686.65
	HSAOR	3.08	92.44	471.17	1511.93
	HSTOR	2.99	93.87	458.45	1527.54
	QSSOR	0.75	19.85	106.87	369.38
	QSAOR	0.73	19.97	108.78	364.51
	QSTOR	*0.66*	*17.80*	*94.22*	*320.61*

The development of path creation was brief, wherein the algorithm plainly picks the lowest temperature value of its adjacent points from the current point. This action remains until the marked target point is achieved. In accordance with the heat transfer analogy with numerical computation, the paths were favourably generated in an obstacle environment as shown in Fig. 3. Each and every single beginning point (green square point) successfully reached the designated destination position (red round point) and evaded various obstacles set in the C-space. Through robot 2D simulator [21], the simulations solely evaluate known static two-dimensional indoor configurations.

To simplify the data, the line graph of the iteration counts, and the time taken for every condition was presented in Figs. 4 and 5, respectively. Clearly shows that all four conditions provide a similar pattern, demonstrating that the QSTOR scheme produced the best outcomes in developing and completing the path as compared to other techniques for both iteration counts as well as CPU time. It can be deduced from the results table and the line chart that utilizing the HS approach has resulted in a nearly and more than 50% reduction than using the standard procedure. Whereas, nearly 75% diminution has taken from QS technique as against conventional technique.

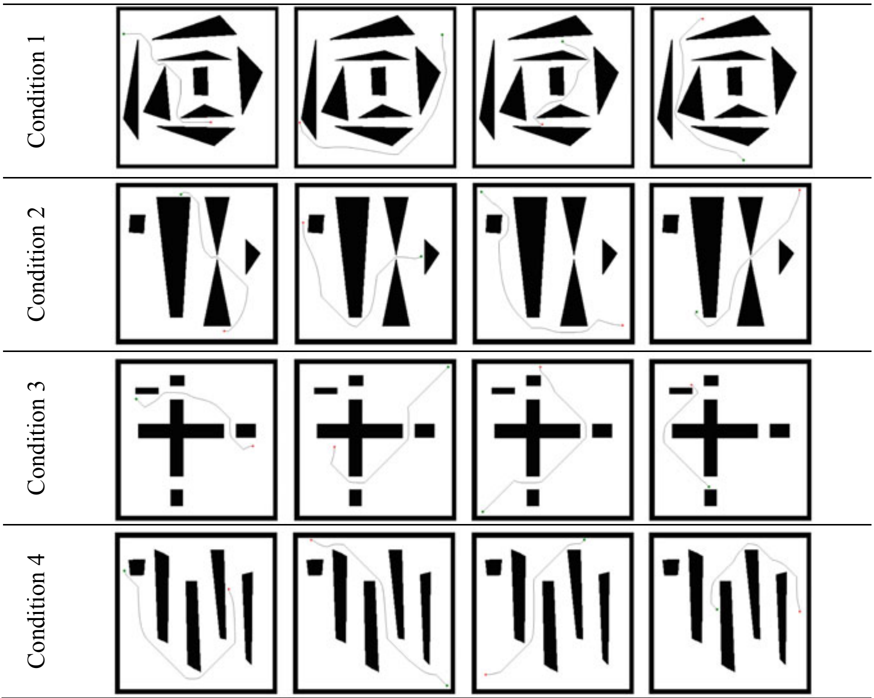


Fig. 3 Produced pathways from various start (green square point) and goal (red round point) points for varied C-space

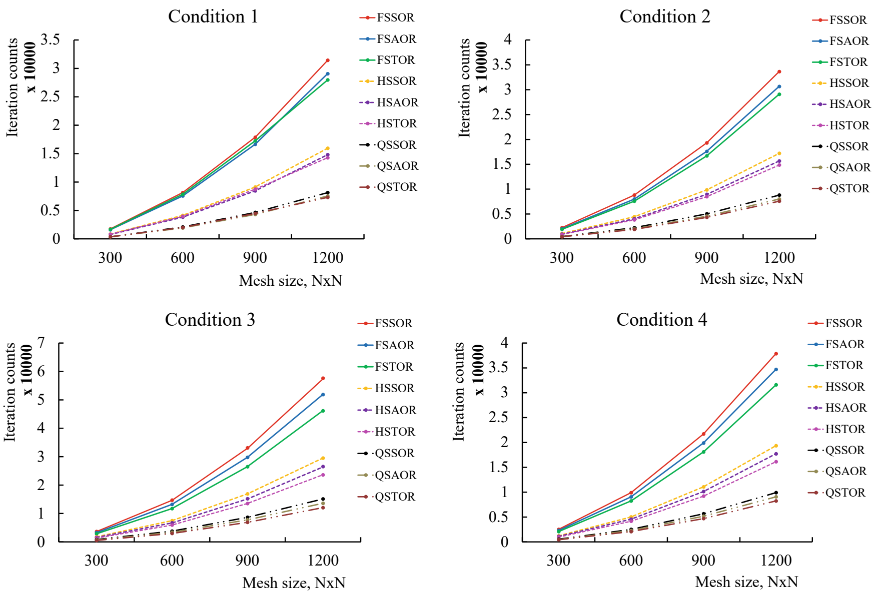


Fig. 4 Performance graph concerning the iteration counts in various C-space sizes

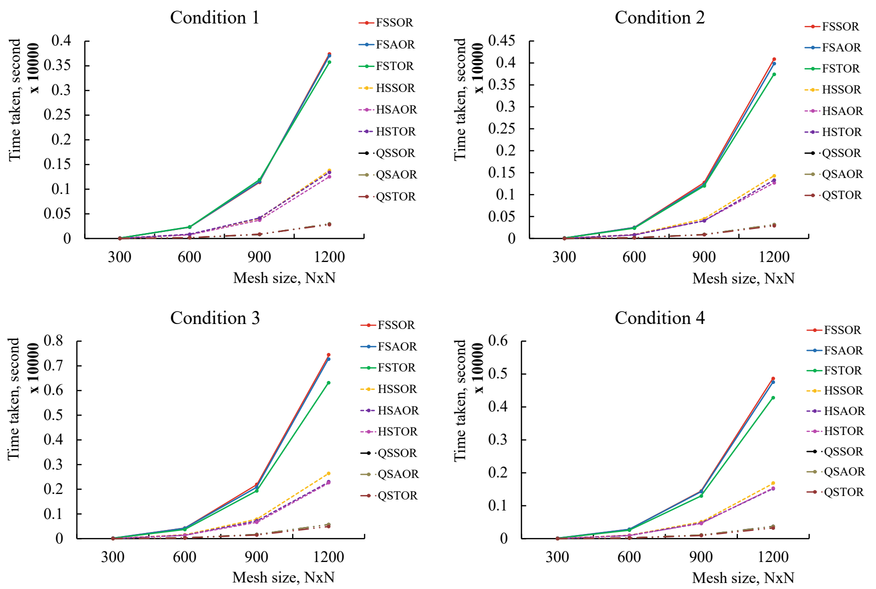


Fig. 5 Performance graph concerning the time taken in various C-space sizes

Concerning the computational complexity analysis of all iterative methods considered, it is assumed that each arithmetic operation requires one unit of computational time. Theoretically, as the complexity analysis is reduced, the number of iterations will become lesser, thus decreasing the CPU time. Even though the number of arithmetic operations for the families of the TOR method is more compared to families of SOR as well as AOR, they converge faster since the presence of weighted parameters [22]. The remaining points, on the other hand, will be omitted in the whole calculation of the computational complexity since they will give no significance to the computation as it does not contribute to the changes in the calculation. After all, the loop for the remaining point is only at one.

It is obvious that the computational complexities of the FS algorithms are reduced drastically by the HS and QS algorithms by approximately 50% and 75%, respectively. As discussed before, only half of the node points are involved during the iteration process of the HS algorithms. For QS algorithms, the iteration process only involves a quarter of node points. Therefore, by reducing the amount of node points involved during the iteration process, convergence can be achieved much faster, thus improving the overall performance of the iterative methods and the path searching process. As for the relation between computational complexities and CPU time, it shows that the higher the complexity, often resulting in higher the CPU time.

5 Conclusions

Owing to the fact the recently developed and newly found techniques, along with the availability of fast machines today, this experiment demonstrates that the solution to mobile path-planning problems through numerical approaches is, in fact, creative and doable. The results table shows that the TOR iterative scheme, in contrast to conventional SOR and AOR techniques, was faster in terms of iteration counts and processing time. Moreover, the implementation of the QS scheme towards the finite difference technique contributes to decreasing computational complexity, leading to the formulation of quarter-sweep two parameter over-relaxation (QSTOR), which provided significant results in this study. The results are unaffected by an increasing number of obstacles because the computing process is only becoming faster as the calculation ignores or disregards the zones occupied by the obstacles. The edge of the proposed algorithm is that it allows the robot to move from starting position to the ending position safely along the shortest path, regardless of the obstacle's size, form, or placement.

Acknowledgements This research was financially supported by Universiti Pertahanan Nasional Malaysia. The authors also acknowledge support from Science Foundation Ireland (SFI) under grant number SFI/16/RC/3918 (confirm), and Marie Skłodowska-Curie grant agreement no. 847577 co-funded by the European Regional Development Fund. The authors declare no conflict of interest and no external data or images were used to support this study.

References

1. Ye Y, Song Z, Zhao J (2022) High-fidelity 3D real-time facial animation using infrared structured light sensing system. Comput Graph 104:46–58
2. Liu Z, Liu Q, Xu W, Wang L, Zhou Z (2022) Robot learning towards smart robotic manufacturing: a review. Robot Comput-Integra Manuf 77:102360
3. Soliman M, Avgeriou P, Li Y (2021) Architectural design decisions that incur technical debt-an industrial case study. Inform Softw Technol 139:106669
4. Ahmad RW, Hasan H, Yaqoob I, Salah K, Jayaraman R, Omar M (2021) Blockchain for aerospace and defense: opportunities and open research challenges. Comput Ind Eng 151:106982
5. Becker M, Faucher P (2021) Recent developments in the implementation of European space surveillance and tracking (EU SST)-security and data policy. J Space Safety Eng 8(2):178–181
6. Dahalan AA, Saudi A, Sulaiman J, Din WRW (2018) Robot navigation in static indoor environment via accelerated iterative method. Adv Sci Lett 24(2):986–989
7. Connolly CI, Burns JB, Weiss R (1990) Path planning using Laplace's equation. In: Proceedings of the IEEE international conference on robotics and automation, vol 3. Cincinnati, OH, pp 2102–2106
8. Connolly CI, Gruppen R (1993) On the applications of harmonic functions to robotics. J Robot Syst 10(7):931–946
9. Al-Khaled K (2005) Numerical solutions of the Laplace's equation. Appl Math Comput 170(2):1271–1283
10. Shivaram KT, Jyothi HR (2021) Finite element approach for numerical integration over family of eight node linear quadrilateral element for solving Laplace equation. Mater Today: Proc 46(9):4336–4340
11. Liu YC, Fan CM, Yeih W, Ku CY, Chu CL (2021) Numerical solutions of two-dimensional Laplace and biharmonic equations by the localized Trefftz method. Comput Mathem with Appl 88:120–134
12. Abdullah AR (1990) The four point explicit decoupled group (EDG) method: a fast Poisson solver. Int J Comput Math 38(1–2):61–70
13. Abdullah AR, Ali NHM (1996) A comparative study of parallel strategies for the solution of elliptic PDEs. Parallel Algorithms and Appl 10:93–103
14. Sulaiman J, Hassan MK, Othman M (2004) The half-sweep iterative alternating decomposition explicit (HSIADE) method for diffusion equation. In: Zhang J, He JH, Fu Y (eds) Computational and information science, LNCS, vol 3314. Springer, Berlin, Heidelberg, pp 57–63
15. Dahalan AA, Saudi A, Sulaiman J, Din WRW (2018) Autonomous navigation in static indoor environment via rotated Laplacian operator. In: AIP conference proceedings 1974:0200352018
16. Dahalan AA, Saudi A (2021) Rotated TOR-5P Laplacian iteration path navigation for obstacle avoidance in stationary indoor simulation. In: iCITES2020, advances in robotics, automation and data analytics, advances in intelligent systems and computing, vol 1350. Springer, Cham, pp 285–295
17. Othman M, Abdullah AR (2000) An efficient four points modified explicit group Poisson solver. Int J Comput Mathem 76:203–217
18. Young DM (1954) Iterative methods for solving partial difference equations of elliptic type. Trans Am Math Soc 76:92–111
19. Ali NHM, Pin FK (2012) Modified explicit group AOR methods in the solution of elliptic equations. Appl Mathem Sci 6(50):2465–2480
20. Hadjidimos A (1978) Accelerated overrelaxation method. Mathem Comput 32(141):149—157
21. Saudi A (2015) Robot path planning using family of SOR iterative methods with laplacian behaviour-based control. PhD thesis, Universiti Malaysia Sabah, Malaysia
22. Kuang J, Ji J (1988) A survey of AOR and TOR methods. J Comput Appl Math 24:3–12

Integrating IoT Sensors to Setup a Digital Twin of a Mixed Model Stochastic System for Real-Time Monitoring

Philane Tshabalala and Rangith B. Kuriakose

Abstract The ongoing digital revolution, commonly referred to as Industry 4.0 has underpinned the importance of real-time monitoring in the manufacturing industry. Real-time monitoring assists with detecting abnormal changes in the production process and also looks for ways to keep the production efficient. An ideal real-time monitoring system must have a sensor network with different types of sensors that will help with data collection. Internet of Things are one of the technologies that can be used for the enabling of real-time monitoring and their ability to transfer the data over a network is automated. The challenge in the manufacturing industry today is having a system that can respond to the abnormal changes in real-time. The existing systems that are currently used for real-time monitoring have some limitations, and therefore this article proposes a digital twin as a possible solution. Digital twins have evolved throughout the years and they will continue to evolve for the next decade, as they play an important role in digital transformation and the vision of smart manufacturing. They have the ability to use live-data as inputs and therefore can predict a number of what-if scenarios, downtime and future faults both in real time or using the historical data. This article discusses all the necessary steps that are needed in developing a data-driven digital twin in the manufacturing industry.

Keywords Digital twin · Internet of things · Real-time monitoring · ThingSpeak · Stochastic systems

P. Tshabalala (✉) · R. B. Kuriakose
Central University of Technology, Bloemfontein, Free State, South Africa
e-mail: philanetshabalala@cut.ac.za

R. B. Kuriakose
e-mail: rkuriako@cut.ac.za

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_24

1 Introduction

The competitiveness of the current global market forces manufacturing firms to stay efficient and produce quality products, as there is no room for errors. This means in order to remain competitive in the market, the firm must be able to adapt to new technologies [1]. One such technology is Real-Time Monitoring (RTM). RTM is critical, as it allows operators and engineers to see and respond to events such as downtime, faults, failures and other issues that might affect the production [2].

Real-time is a term that can be defined as the ability of the system to respond to any swift change in a way that the response is almost at the same time that the change/event occurred on [2]. Real-time data collection is challenging as it must not interfere with the running application [3]. For this reason, the simulation models (such as digital shadows [4] and optimal solutions [5]) that are currently available do not take in real-time data as inputs for simulations or tests, and therefore these models have a number of limitations [4].

The disadvantage of these traditional models is that they are not connected to the physical system which means there is no real-time data collection and therefore not a solution to the challenges in today's market [4]. Some of the challenges in today's market, but not limited to, are high demands in product variety, shorter lead times, real-time monitoring and controlling. This can assist in pre-predicting production flaws and the design reconfigurations [6].

Digital twins (DTs) are part of the emerging fourth industrial revolution and they are known for their ability of linking the physical and digital worlds [7], therefore making them a possible solution to a number of challenges in today's global market. A DT can be defined as a mirror image of a real-world object presented in a digital world, for real-time monitoring, controlling, simulation and testing [8].

This article looks at how to develop a data-driven DT in the manufacturing industry, from the selection of different Internet of Things (IoT) sensors to having a real-time monitoring model. The article is divided into five sections. Section 2 discusses mixed model stochastic systems, available methods for real-time monitoring, IoT sensors and ThingSpeak. Section 3 discusses the aim and methodology of this research. Section 4 looks at the results. Section 5 discusses the conclusion and future works.

2 Background

2.1 Mixed Model Stochastic System

Assembly lines are an integral part of a manufacturing process and are used to move products from one workstation to the next [9]. The time a workstation takes to complete the tasks assigned to it is referred to as task time [10]. By nature of operation an assembly line can either be deterministic or stochastic [9]. A deterministic system is a system where the inputs are known or pre-determined while a stochastic system is a system in which the inputs are not pre-determined [4].

Stochastic systems fit well with today's market, as one of the demands of the fourth industrial revolution (4IR) is the need of customized products [11]. Hence, a mixed model stochastic (MMS) system refers to a system that is designed or developed to produce a variety of products using different inputs [4]. This article discusses the setup and integration of IoT sensors to a MMS system with the aim of creating a data-driven digital twin that will showcase the importance of real-time monitoring in smart manufacturing.

2.2 Current Methods That Are Used for Real-Time Monitoring

RTM helps with identifying the actual time a fault occurred on and the time it was attended to and resolved, by sending the user alerts and notifications [3]. These are some of the methods or techniques that are available and currently used for real-time monitoring across the globe:

- Data Acquisition (DAQ) System

DAQ can be defined as a technique that is used for digitalizing the data from different sensors so it can be stored in computers, where it can be displayed and analyzed. DAQ system are used in various industries for tests, measurements and automation, and are known to be the best in measuring the current and voltage signals [12]. A DAQ system consists of and not limited to sensors, communication links, signal processors, computers, DAQ software, just to name a few [12].

- Lean Manufacturing (LM)

Lean manufacturing can be defined as an approach that aims at exceeding the customer's expectations by continuously reducing all kinds of waste (overproduction and over processing) in the manufacturing process to enhance the production and produce products at lower costs [13, 14].

- Cloud Manufacturing (CM)

Cloud manufacturing is a new technology that was developed to transform the manufacturing industry [15], it can be defined as a new paradigm that is made of the existing manufacturing models with the support of new technologies such as IoTs and cloud computing with the aim to keep the production efficient and promote collaborations within the industry [16]. CM makes use of IoT technologies such as radio frequency identification (RFID) and wireless communications for real-time monitoring [17].

However, these techniques have their limitations. The limitation of DAQ systems is that they come with built-in sensors and therefore can only be utilized for that specific application and also initial training is required for new users and programmers [18].

The limitation of LM is that it takes time to give feedback (meaning the response is not in "real-time"), as it wants to give out an accurate response. LM must be used with other emerging technologies such as cloud manufacturing in order to defeat this time factor [1]. The limitations of CM are network outages and the risk of unexpected downtime since CM is an Internet-based system [19].

2.3 IoT Sensors

IoTs are one of the new technologies that were developed in the emerging Industry 4.0, which can be defined as the technology that enables interaction between the physical and the digital worlds with the goal of sharing resources, data and information [20]. The selection of IoT sensors is one of the first and important stages when setting up a real-time monitoring system, as these sensors enable the system to collect data and detect the type of data coming from the physical object [21]. Selection of wrong sensors may result into the developed system not functioning as expected [21].

2.4 ThingSpeak

The advent of IoTs has led to the release of hundreds of different platforms, one being ThingSpeak [22]. ThingSpeak allows collection, virtualization and analyzation of live data and has built-in libraries to support IoT devices [22]. ioBridge developed this open source application back in 2010 [23]. The easy interface and data processing is what makes ThingSpeak stand out from the other IoT platforms, credit can be given to the support of the MATLAB language and MATLAB toolboxes. The Hypertext Transfer Protocol (HTTP) and the Message Queuing Telemetry Transport (MQTT) are the two communication protocols used by ThingSpeak to prove APIs [22]. Thingspeak's main component is the channel, which stores the data sent from various IoT devices [23].

3 Methodology

The selection of different IoT sensors shown in Fig. 1 was done based on the aspects that this system wishes to monitor. Arduino uno is a microcontroller that functions as the brain of this system, as it is the one responsible for the collection of data from the IoT sensors. The data must be sent to the cloud (ThingSpeak) after collection and this system will do that with the help of the ESP-01 (a Wi-Fi module), as it enables connection between the microcontroller and the Wi-Fi network.

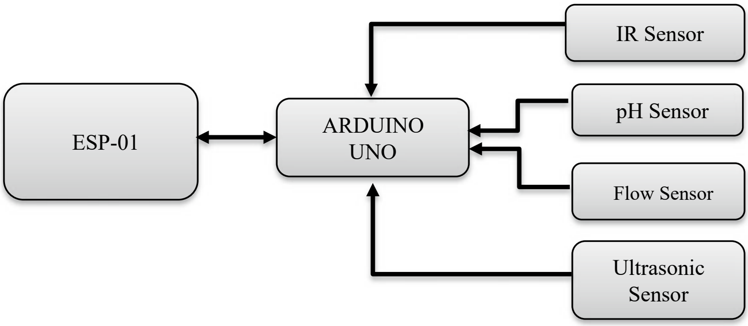


Fig. 1 IoT sensors setup block diagram

This article uses the case study of the smart water bottling plant at the Central University of Technology, Free State. The water bottling plant is made of three smart manufacturing units (also known as SMUs) which are driven by Programmable Logic Controllers (PLCs), for filling, capping and packaging. Figure 2 [24] shows the proposed setup and how the different IoT sensors will be installed to monitor the raw materials in these SMUs by creating a digital twin of this water bottling plant.

The digital twin of the smart water bottling plant [25] was developed using MATLAB/SIMULINK. This model is capable of taking in real-time data from the IoT sensors (sent to ThingSpeak) and use it as inputs to monitor in real time the raw materials (Fig. 3).

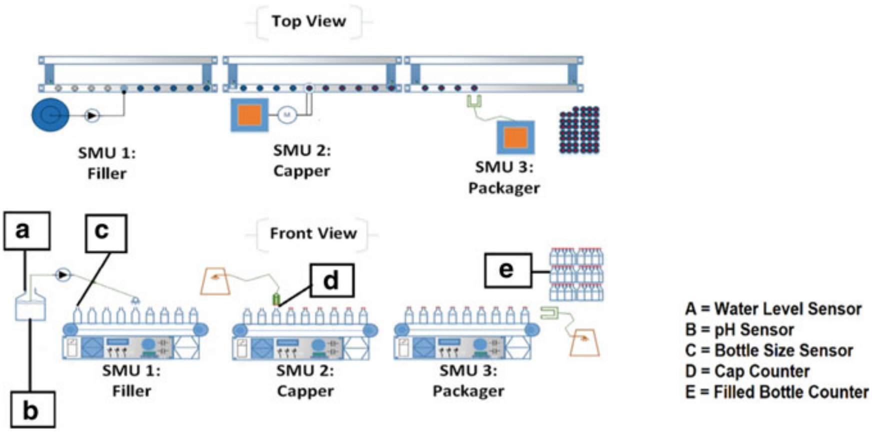


Fig. 2 Proposed experimental setup for the digital twin

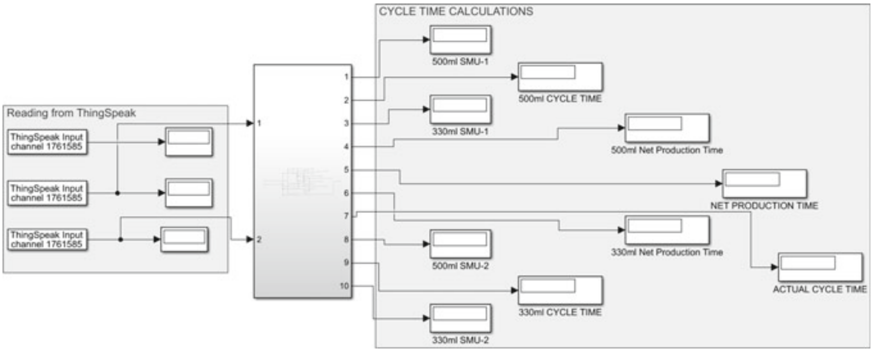


Fig. 3 Digital twin model on SIMULINK

4 Results

As discussed in the previous section, the microcontroller collects the live data from the selected IoT sensors and send it to ThingSpeak for visualization and data analysis. The data is split to different fields of the ThingSpeak channel and presented in a graphical format, as shown in Fig. 4.

A channel with the name "WATER BOTTLING PLANT" was created on ThingSpeak. The channel consists of three fields for real-time monitoring of the water level, number of bottles and the water flowrate. Figure 4 shows the live data coming from

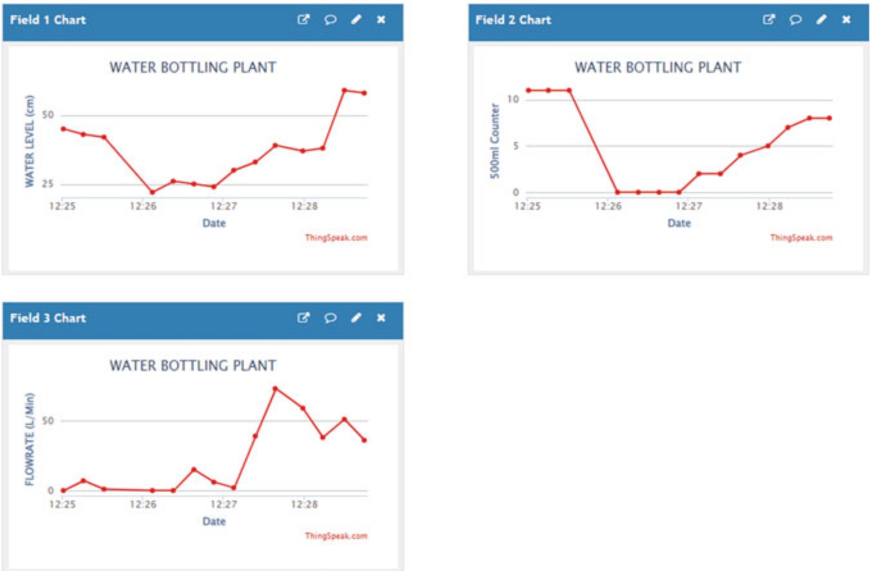


Fig. 4 ThingSpeak channel "WATER BOTTLING PLANT"

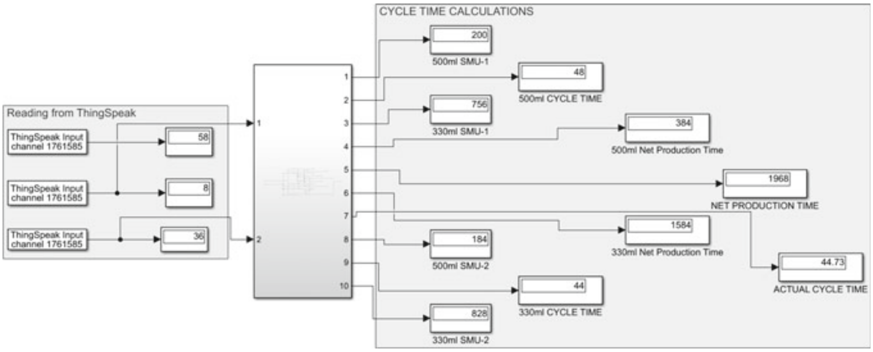
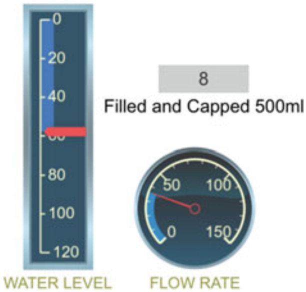


Fig. 5 Digital twin taking live data as inputs

Fig. 6 Dashboard



the three IoT sensors. It should be noted that data displayed was for testing if the sensors can be integrated to send data to the channel fields in real time.

The developed digital twin in Fig. 5 shows how the live data will be streamed and used as inputs to enable real-time monitoring and calculations of cycle time. Figure 6 shows how the data will then be shown on the dashboard (on SIMULINK) for the user to have a simplified visualization.

5 Conclusion and Future Works

This article focused on the integration of IoT sensors for setting up a data-driven digital twin, so as part of future work more features will be added to the digital twin and it will be used for predicting possible bottlenecks, calculating the cycle time and to minimize the downtime. A digital shadow will also be created for the water bottling plant and the results will be compared with that of the digital twin. It is hypothesized that the ability of the digital twin to use live data as inputs will reduce the cycle time better compared with the digital shadow.

References

1. Kumar M, Vaishya R, Parag (2018) Real-time monitoring system to lean manufacturing. Proc Manuf 20:135–140. https://doi.org/10.1016/j.promfg.2018.02.019
2. Kebande VR, Karie NM, Ikuesan RA (2021) Real-time monitoring as a supplementary security component of vigilantism in modern network environments. Int J Inf Technol 13(1):5–17. https://doi.org/10.1007/s41870-020-00585-8
3. Mahadevan M (2012) Data collection and performance monitoring of real-time parallel data collection and performance monitoring of real-time parallel systems systems. vol 76. pp 2012–76. [Online]. Available: https://openscholarship.wustl.edu/cse_research/90
4. Tshabalala P, Kuriakose RB (2022) Analyzing the performance of a digital shadow for a mixed-model stochastic system. In: Sharma H (ed) Lecturer notes in networks systems, Singapore, Springer
5. Kuriakose RB, Vermaak HJ (2018) A review of the literature on assembly line balancing problems, the methods used to meet these challenges and the future scope of study. Adv Sci Lett 24(11):8846–8850. https://doi.org/10.1166/asl.2018.12359
6. Helu M, Morris K, Jung K, Lyons K, Leong S (2015) Identifying performance assurance challenges for smart manufacturing. Manuf Lett 6:1–4. https://doi.org/10.1016/j.mfglet.2015.11.001
7. Polini W, Corrado A (2020) Digital twin of composite assembly manufacturing process. Int J Prod Res 58(17):5238–5252. https://doi.org/10.1080/00207543.2020.1714091
8. Melesse TY, Di Pasquale V, Riemma S (2021) Digital twin models in industrial operations: state-of-the-art and future research directions. IET Collab Intell Manuf 3(1):37–47. https://doi.org/10.1049/cim2.12010
9. Kuriakose R et al. (2019) Optimization of a real time web enabled mixed model stochastic assembly line to reduce production time. 2019, [Online]. Available: https://doi.org/10.1007/978-981-15-0108-1_5
10. Kuriakose RB, Vermaak HJ (2020) Customized mixed model stochastic assembly line modelling using simulink. pp 2–7. https://doi.org/10.5013/IJSSST.a.20.S1.06
11. Renard P, Alcolea A, Ginsbourger D (2013) Stochastic versus deterministic approaches, November 2018
12. Sarma P, Singh HK, Bezboruah T (2018) A real-time data acquisition system for monitoring sensor data. Int J Comput Sci Eng 6(6):539–542. https://doi.org/10.26438/ijcse/v6i6.539542
13. Bhamu J, Sangwan KS (2014) Lean manufacturing: literature review and research issues. Int J Oper Prod Manag 34(7):876–940. https://doi.org/10.1108/IJOPM-08-2012-0315
14. Gobinath S, Elangovan D, Dharmalingam S (2015) Lean manufacturing issues and challenges in manufacturing process– a review. Int J ChemTech Res 8(1):45–51
15. Rahman MNA, Medjahed B, Orady E, Muhamad MR, Abdullah R, Jaya ASM (2018) A review of cloud manufacturing: issues and opportunities. J Adv Manuf Technol 12(1):61–76
16. Ren L, Zhang L, Tao F, Zhao C, Chai X, Zhao X (2015) Cloud manufacturing: from concept to practice. Enterp Inf Syst 9(2):186–209. https://doi.org/10.1080/17517575.2013.839055
17. Zhong RY, Wang L, Xu X (2017) An IoT-enabled real-time machine status monitoring approach for cloud manufacturing. Proc CIRP 63:709–714. https://doi.org/10.1016/j.procir.2017.03.349
18. Abidin AFZ, Jusoh MH, James E, Al Junid SAM, Mohd Yassin AI (2015) Real-time remote monitoring with data acquisition system. In: IOP conference series material science engineering, vol 99(1). https://doi.org/10.1088/1757-899X/99/1/012011
19. Viswanathan P (2018) Pros and Cons of Cloud Computing. Int J Sci Res 5(7):2013–2016
20. Rana A, Kumar A (2021) A review paper on internet of things (IOT). Asian J Multidimens Res 10(10):166–172. https://doi.org/10.5958/2278-4853.2021.00915.0
21. Hirayama M (2016) Sensor selection method for IoT systems—focusing on embedded system requirements. In: MATEC web conference, vol 59. https://doi.org/10.1051/matecconf/20165901002

22. De De Nardis L, Caso G, Di Benedetto MG (2019) ThingsLocate: a ThingSpeak-based indoor positioning platform for academic research on location-aware internet of things. Technologies 7(3):50. https://doi.org/10.3390/technologies7030050
23. Nettikadan D, Raj S (2018) Smart community monitoring system using thingspeak IoT plaform. Int J Appl Eng Res 13:13402–13408. [Online]. Available: http://www.ripublication.com
24. Gericke G, Kuriakose R, Vermaak H (2019) Design of digital twins for optimization of a water bottling plant
25. Kuriakose RB, Vermaak HJ (2020) Designing a simulink model for a mixed model stochastic assembly line : a case study using a water bottling plant. J Discret Math Sci Cryptogr 23(2):329–336. https://doi.org/10.1080/09720529.2020.1741184

Deep Learning-Based Multi-task Approach for Neuronal Cells Classification and Segmentation

Alaoui Belghiti Khaoula, Mikram Mounia, Rhanoui Maryem, and Yousfi Siham

Abstract Neurodegenerative diseases are causing death increasingly due to lack of treatment, and the challenges professionals in the medical field are having regarding the precision of the diagnosis, which is extremely hard for an individual to accurately determine the progression of the disease based on cell images. As in many other issues deep learning has helped optimize the diagnosis process and automize the cell lines classification and segmentation; we used a multi-task architecture to address the two tasks by one accurate model to determine the cell type and extract a precise segmentation of the cells, facilitating the diagnosis process for the professionals in medical field, and refer to the neurodegenerative disease that should be treated. Therefore, to achieve accurate results we have based our solution's architecture on state-of-the-art segmentation model U-Net for medical images with pre-trained classification models (VGG16 and MobileNetv2 separately) as backbones for classification task. We applied our models on the Sartorius cell instance segmentation dataset containing phase-contrast microscopy (PCM) images of human neuronal cells along with their annotations; despite the small number of images provided, the dataset contains a high amount of annotated cells. Combining the two tasks in one model reached a segmentation performance of 79.6 and 80.1% for the U-Net model with MobileNetv2 backbone and VGG16, respectively, outperforming the single task U-Net that only achieved 75%, simultaneously a classification accuracy of 99.6 and 99.7% with the multi-task models with the VGG16 and the MobileNetv2 backbone compared to 95.7 and 93%, respectively, for the VGG16 and the MobileNetv2 classification models,

A. B. Khaoula (✉) · M. Mounia · R. Maryem · Y. Siham
Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco
e-mail: khaoula.alaoui-belghiti@esi.ac.ma

M. Mounia
e-mail: mmikram@esi.ac.ma

R. Maryem
e-mail: mrhanoui@esi.ac.ma

Y. Siham
e-mail: syousfi@esi.ac.ma

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_25

proving that this approach gets a high accuracy in terms of classification and segmentation tasks and outperforms the mono-task models applied in the context of the Sartorius cell competition.

Keywords Neuronal cells · Segmentation · Classification · Deep learning · Medical imaging · Multi-task

1 Introduction

Neurodegenerative diseases such as Alzheimer's and Parkinson's [11] are causing death increasingly due to the absence of cures and the basic symptom-based treatments, leading to more intention on early and precise diagnosis of such diseases. Professionals in the medical field are having several challenges regarding the precision and speed of neuronal diseases diagnosis, which is difficult for a human being to accurately determine the regression of neuronal cells. The process of analyzing the types and morphology of neuronal cells have been a revolution in neuroscience [7]. Microscopy imaging techniques, especially phase contrast microscopy (PCM) [1], are widely used to capture cells morphology, making the images acquisition an interesting boost for computer-aided analysis of these cells structure and behavior. As in many other fields, machine learning and especially deep learning advanced techniques, either through supervised or unsupervised learning [4, 14], interfere to optimize the process; in the medical field it helps in several diagnosis process [8] and disease detection [15], in our case intending to automate accurately the neuronal cells analysis. Recent image processing techniques have made it easier to capture in a detailed level as cell lines and analyze their morphology using computer vision (CV) tasks, mainly the classification and segmentation tasks, considering how important is the segmentation task when made accurately with computer-aided methods in order to monitor the cell's pattern and then the progression of the studied neuronal disease. Current cell analysis solutions have limited the model use in one main task either classification of cell lines or segmentation, instance to facilitate the cell counting or semantic to focus on neuronal cells morphology even though there are several datasets that used professional's big efforts to present cells labels and masks in order to facilitate the classification as the segmentation tasks. A previous study in 2021 presented an interesting large cell segmentation dataset called LIVECell [5], with more than 5K PCM cell images representing eight different cell lines. Several deep learning solutions validated on the LIVECell dataset under-represented the neuronal cell line Shsy5y with low segmentation accuracy, which may go back to the special morphology of this cell type making the segmentation process extremely challenging. Regarding the morphology and other challenges that faces neuronal cell image segmentation, such as the background noise and contrast of cell boundaries, we notice multiple advances using several applications of deep learning models on cell lines analysis. Models based on deep neural networks have been applied to a wide range of computer vision (CV) tasks, mainly the classification [12], and the semantic segmentation tasks [16] However, existing studies have not dealt with the segmentation as well as the classification task at the same time with effective solutions; most of the proposed solutions are based on mono-task models, which are directly effected with narrow data for extensive model training. To address this issue, we choose to base our solution on multi-task learning [2] to implement a effective architecture based on the close tasks of classification and segmentation sharing many characteristics which helps creating one backbone and differentiate the head of the model to precisely apply the two tasks, in order to optimize the material and time resources consumption on relatively small datasets as provided for the neuronal cells analysis. In order to achieve such high precision results, our deep learning approach is based on convolutional neural networks (CNNs) multi-task models able to classify and segment cell images at the same time. As a backbone, we used the state of the art, medical images segmentation, U-Net model [17] and the pre-trained classification models, VGG16 [19] and MobileNetv2 [18] for the encoder. We evaluated the proposed architectures on the Sartorius cell instance segmentation dataset recently uploaded in a competition context that have shown better performance than segmentation only with U-Net, or classification with VGG16 or MobileNetv2 .

To describe the solution, we first review some related works to the techniques and the dataset used in Sect. 2, and then, we cover in Sect. 3 the dataset with a detailed description of the proposed architecture. After that in Sect. 4, we list the experimental results along with discussion. Lastly, in Sect. 5 we summarize the paper and point out future research directions.

2 Related Works

Neuronal cells classification and segmentation is extremely important to ensure a precise monitoring of neurological disorders, several works have addressed the subject using different datasets; made with professionals efforts to provide a computer treatable data as in recent LIVECell dataset and its follow Sartorius Dataset, or in deep learning solutions that employ deep neural network (DNN)-based models for cells segmentation. [21] presented an hierarchical neural network applying the object detection and segmentation tasks, with full use of features at different levels. Another work of [13] used auxiliary cells in an attempt to over-parameterize the semantic segmentation model's architecture, for [22] focuses its solution on a box-based cell instance segmentation using keypoints detection, providing a cell detection with bounding boxes as well as segmentation masks. Yi et al. [23] added an attention mechanism to a special architecture merging a single-shot multi-box detection and segmentation using U-Net model. In another approach, [9] proposes a solution for cells tracking multi-task model based on the classification and the detection with bounding boxes; with the same idea [3] approaches the problem from a time tracking view; in a weekly bases it combines the detection and segmentation tasks in an intensive multi-task model architecture.

Therefore, we notice that our solution differs from the existing DNN-based cell images analysis architectures in several levels. First, the use of newly provided datasets which have benefited from the advances of technology is to provide precise segmentations and classifications of the studied cells, and also, the effect of transfer learning for cell-type classification has not been extensively used. Given the similarity between the classification and segmentation of medical images, we cannot skip how useful it is to combine the two tasks in one model and reduce the resources consumption in a relatively small duration with the small dataset provided. These common but interesting techniques helped us to provide a better performing solution for the neuronal cells classification and segmentation through several experiments and comparisons.

3 Materials and Methods

3.1 Dataset Description

The dataset we used is the Sartorius Cell Instance Segmentation dataset (SCIS),[1] generated for a Kaggle competition recently finished on the December 30, 2021. SCIS dataset is sort of a follow-up dataset of the large LIVECell dataset, annotated manually by professionals and validated by medical experts. It represents eight different cell lines, including Shsy5y, with a unique neuronal morphology and overlapping cells affected its segmentation accuracy. So they developed the new SCIS dataset focusing the image analysis more on neuronal cells segmentation.

The SCIS dataset consists of three neuronal cell types, namely Cort, Shsy5y, and Astro. The dataset consists of a total of 606 PCM image samples of all three types, including 320 Cortical neurons (Cort), 155 Shsy5y, and 131 astrocytes (Astro) samples. Shsy5y cells may be transformed to various types of functioning neurons by adding particular substances, making it a model for neurodegenerative illnesses. In addition, the Shsy5y cell line has been widely employed in experimental neurological investigations, including analysis of neuronal development, metabolism, and function in relation to neurodegenerative processes, neurotoxicity, and neuroprotection [10]. Astrocytes are a type of glial cell that outnumber neurons by a factor of five. They tile the entire central nervous system (CNS) and perform a variety of important complicated tasks in a healthy CNS [20].

From the SCIS samples, the average number of annotations per image is 34 Cort, 337 Shsy5y, and 80 Astro cells. For the mask area represented by number of pixels is 240 for Cort, 224 Shsy5y and 906 Astro. As noticed, the Astro cells are highly represented than the other cell lines, while Shsy5y cells show the highest density. All cell images have similar dimensions, 520×704. The SCIS dataset contains PCM

[1] https://www.kaggle.com/c/sartorius-cell-instance-segmentation.

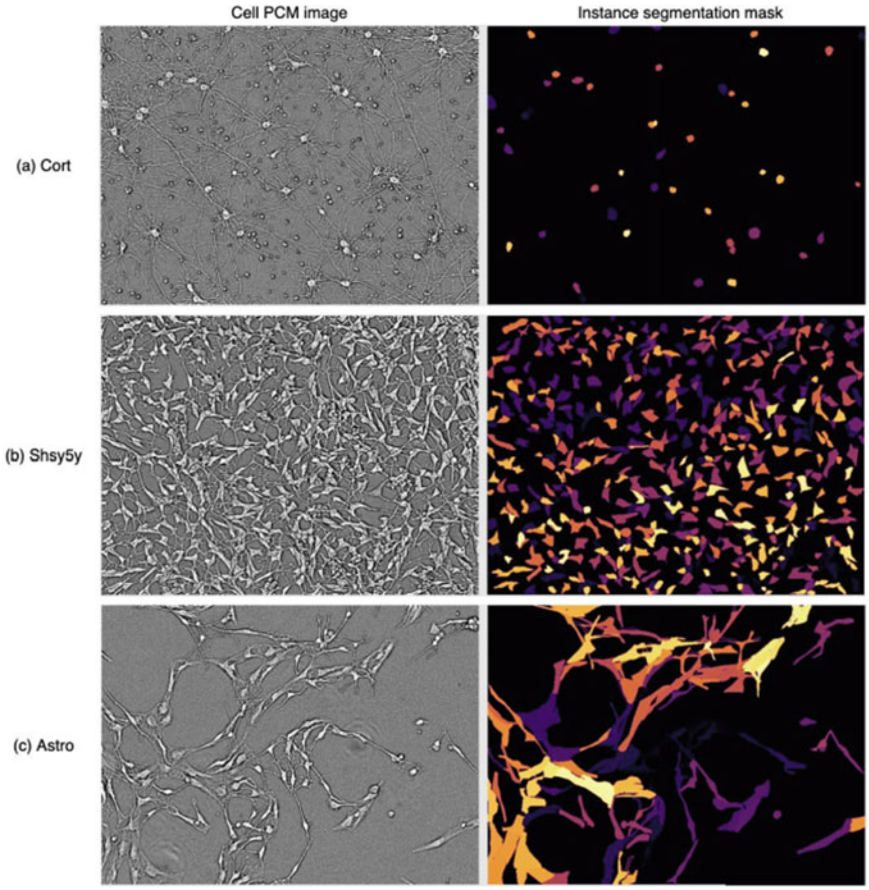
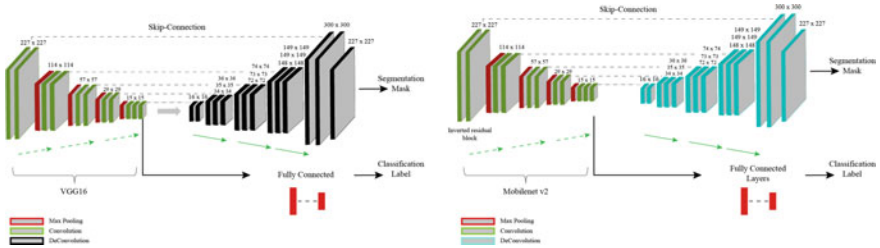


Fig. 1 SCIS cell types samples: for each cell line we present in left the original cell PCM image and in the right side its segmentation mask

images each one is related to several mask annotations. In Fig. 1, there are three different samples from the three cell lines with an original PCM image in the left side and the human-annotated mask in the right side, the first one represents the cort cell line, followed by Shsy5y type and last Astro cell, which represents the difference between the three cell types in terms of shape, size, and density.



(a) With VGG16 (b) With MobileNetv2

Fig. 2 Multi-task U-Net architecture

3.2 Multi-task Learning

Benefiting from the humans way of thinking and processing learning new things, mostly based on the knowledge gained from previous activities; the multi-task learning (MTL) [2] comes as an implementation of such a robust process to use the relativity between two tasks in order to apply them simultaneously and increase their performance, which makes it a more efficient solution in limited data cases. MTL can learn robust representations between related tasks. These shared representations increase the efficiency of the data, which leads to better performance and a mitigation of the risk of excessive overfitting.

In this article, we use the multi-task architecture based on U-Net; considering as it was designed specifically to handle the segmentation of complex medical images and has proven its good performance in the medical field as well as in many others. Using hard parameter sharing, we combined transfer learning and multi-task learning [6], where the decoding block uses pre-trained models, VGG16 (Fig. 2a) and MobileNetv2 (Fig. 2b) on the extensive ImageNet Dataset. The use of transfer learning allows for robust feature learning and reduces the number of parameters to be trained. It also allows the model to converge much faster compared to a model trained from scratch.

Thanks to the special architecture of the multi-task network, the components are able to inter-share the features extracted from previous layers, so that they can be used for prediction of both tasks on the neural cells. The architecture Fig. 2a is based on an encoder–decoder network that uses skip-connections to overcome the information loss through the multiple layers.

The encoder part is utilized in order to extract the image perspective, using a classical stack of convolution layers and Maxpool, adding Fully Connected Layers to generate a label related to the inserted image. Bottleneck comprises a compressed representation of the input data; we are then able to generate a cell line as a classification label and a mask containing the segmented cells.

3.3 Training and Evaluation Metrics

In this section, we provide additional information on the training process and the metrics used to evaluates our models performance. While the training process of the multi-task models took an optimal time considering the combined tasks of classification and segmentation of the cell images, we first preprocessed the CSIS dataset images.

Dataset Preprocessing: As a first step, we needed to extract the masks images from the annotations provided in the CSV file of the training set, with overlapping the cells labels in order to provide a precise mask of the cells which later facilitates the feature extraction for our segmentation task. For an efficient learning, we used also data augmentation (DA) techniques to overcome the data sparsity and effectively improve the learning process by extending our training samples with multiple transformations, using rotation, shearing, and flipping techniques.

Accuracy: In case of binary classification, the model tries to calculate whether the samples are Positive P or Negative N. The model's response may be listed as a confusion matrix, which is mostly separated into four different groups: Positive samples properly classified which we call the $(TruePositivesTP)$. The samples misclassified as positives are named $(FalsePositivesFP)$. Then, we call the negative samples that were properly identified $(TrueNegativesTN)$. Similarly, the $(FalseNegativesFN)$ are the positive examples predicted as negative. Based on these values, we are able to predict the model's accuracy, by calculation the proportion of correct samples TP to the total number of predictions.

Dice Coefficient: A widely applicable metric in several fields as computer vision and Natural Language Processing (NLP), in most cases the Dice Similarity Coefficient (DSC) is used to evaluate the segmentation task performance measuring the similarity between the segmented pixels and the ground truth. The coefficient value goes from 0 to 1, no overlap of the pixels to complete overlap.

Loss Function: We used a linear combination of task losses; for the multi-class classification task we based on the categorical cross-entropy loss as in the equation:

$$\text{loss}_C(P, G) = \frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}(G_{ic}\log(P_{ic}))$$

where P is the vector combining the predicted probability of each class, comparing it to G as grading ground truth. N represents the examples of training samples and C total of categories, As output P_{ic} represents the probability distribution for ith observation referring to the c class, so the target G_{ic} may be considered true.

For the semantic segmentation task, we get as output a prediction map containing a class label as integer in each pixel. Considering its value as binary classification to determine whether the studied pixel belongs to the cell or to the image background. Therefore, the loss equation represents a binary cross-entropy composed by the average loss of all examples, considering n total examples, y as true label, y_i an element of the label, $\hat{y}$ the model's prediction and $\hat{y}_i$ an element of the prediction:

$$\text{loss}_S(y, \hat{y}) = -\frac{1}{n}\sum_{i=0}^{n}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

Finally, the multi-task model's loss function represented by $Total\_loss$ is a linear combination of the classification and segmentation loss.

In the coming part, we discuss the results of the implementation of our multi-task architecture regarding the segmentation and the classification tasks.

4 Results and Discussion

	Mono-task		Multi-task	
	MobileNetv2 (%)	VGG 16 (%)	MobileNet encoder (%)	VGG16 encoder (%)
Classification	95.7	93	**99.7**	99.6
Segmentation	75		79.6	**80.1**

Neuronal cells importance: As much as it is hard to precisely overcome the individual diagnosis, presenting a computer use ready datasets is extremely heard to segment and label, which proves how important the neuronal cell datasets even with a small number of samples; all in the purpose to facilitate the diagnosis process for professionals in health sector. In order to assist the development of new cures for the deadly character and the risks that presents neurological disorders. Despite having a relatively small number of samples in the SCIS dataset used, we have achieved interesting classification and segmentation performance using recent deep learning techniques, using the similarity between the two tasks as the previous learning on large datasets to classify the cell lines.

We noticed that the multi-task models either with VGG16 encoder or MobileNetv2 performs highly compared to mono-task U-Net model for the segmentation task and VGG16 or MobileNetv2 for the classification task. Regarding the classification task, the multi-task model MobileNetv2 and VGG16 performed relatively similar with an accuracy of **99.7%** and **99.6%**, respectively. Compared to single task classifiers, the multi-task one improves accuracy of classification from **93%** to **95.7%**, respectively, for the mono-task VGG16 and MobileNetv2, up to **99.7%** for MobileNetv2 multi-task classification model.

Regarding the segmentation task, the multi-task model produced the best results on three cell lines, we have got a dice coefficient of **79.6%** and **80.1%**, respectively,

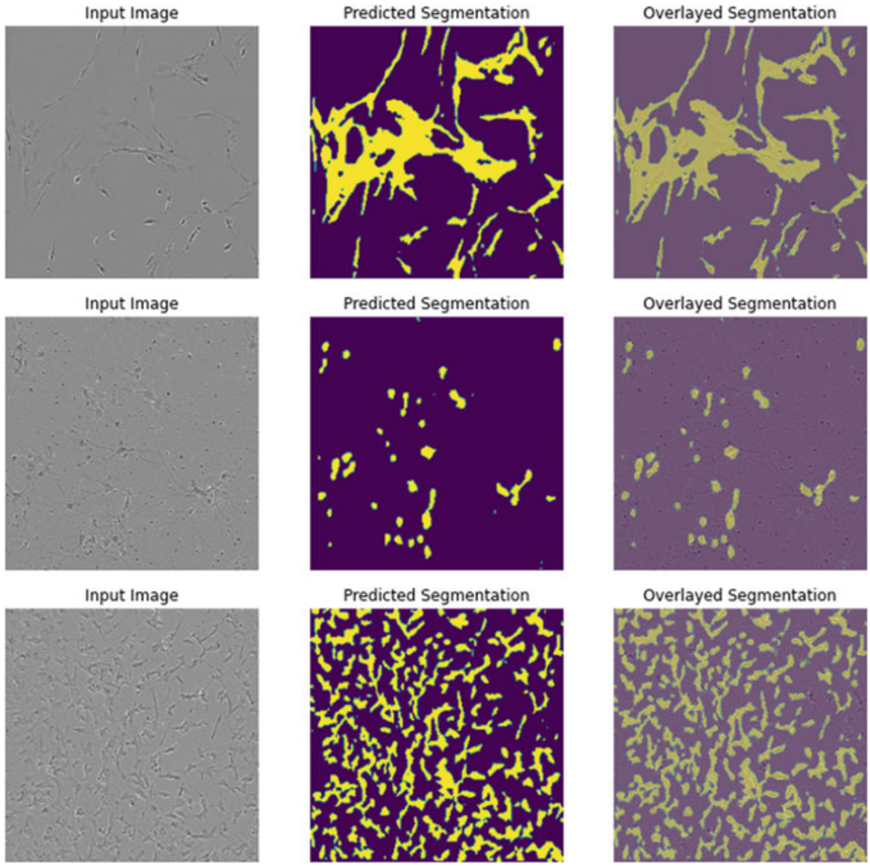


Fig. 3 Segmentation results

for the MobileNetv2 backbone and VGG16 against **75%** for the mono-task U-Net model. Our model is able to accurately segment the cells in the PCM images as it has also determined with high accuracy that the first represented image id astro cell line followed by a cort cells and then the segmentation of Shsy5y cell type, Fig. 3.

As noticed previously, our approach outperforms the mono-task models in all different tasks analyzed, also that the multi-task model with VGG16 encoder outperforms the model with MobileNetv2 backbone in terms of segmentation as complex task who requires a heavy computing and takes longer training time, but in terms of classification, the light architecture of MobileNetv2 helped the model outperform the VGG16 backbone model. Although both models proved to be an optimal solution to facilitate the diagnosis process as well as the progression analysis of the cells types and morphology.

5 Conclusion

Deep learning-based neuronal cell images analysis has been an interesting pathway that can help uncover new cures in neuroscience. Current approaches fall short in segmentation precision basically due to the data scarcity, irregular morphology of cell lines, low quality, and cells overlapping. Despite these challenges, we proposed a multi-task learning approach to cover the classification task along with the semantic segmentation in order to treat each cell line separately and improved performance over state-of-the-art single-tasking models.

In this approach, we tried to focus on the main challenge commonly encountered in computer-aided neuronal cells analysis, namely the lack of annotated images by the multi-task architecture and dealing with two tasks executing one model. The results show that the approach is more effective than the single-tasking techniques, and it could be applied to other datasets to meet different medical problems and prove its efficiency to optimize several diagnostic processes.

Although these solutions prove a better performance than other models dealing with the classification and semantic segmentation issues, the dataset competition was provided in an instance segmentation context; we centered our attention on the cells morphology rather than cell counting that is possible only with instance segmentation which we suggest future search direction of a model architecture using an instance segmentation as Mask-RCNN to uncover more detailed cellular mechanisms and get to a robust solution where we are able to differentiate not only between cell lines but also between cells of the same type.

References

1. Burch C, Stock J (1942) Phase-contrast microscopy. J Sci Inst 19(5):71
2. Caruana R (1997) Multitask learning. Mach Learn 28(1):41–75
3. Chamanzar A, Nie Y (2020) Weakly supervised multi-task learning for cell detection and segmentation. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI). IEEE, pp 513–516
4. Chtouki K, Rhanoui M, Mikram M, Yousfi S, Amazian K (2022) Supervised machine learning for breast cancer risk factors analysis and survival prediction. In: 6th International conference on big data and internet of things
5. Edlund C, Jackson TR, Khalid N, Bevan N, Dale T, Dengel A, Ahmed S, Trygg J, Sjögren R (2021) Livecell-a large-scale dataset for label-free live cell segmentation. Nat Methods 18(9):1038–1045
6. Foo A, Hsu W, Lee ML, Lim G, Wong TY (2020) Multi-task learning for diabetic retinopathy grading and lesion segmentation. In: Proceedings of the AAAI conference on artificial intelligence vol 34, pp 13267–13272
7. Halfter W, Chiquet-Ehrismann R, Tucker RP (1989) The effect of tenascin and embryonic basal lamina on the behavior and morphology of neural crest cells in vitro. Dev Biol 132(1):14–25
8. Harnoune A, Rhanoui M, Mikram M, Yousfi S, Elkaimbillah Z, El Asri B (2021) Bert based clinical knowledge extraction for biomedical knowledge graph construction and analysis. Comput Methods Programs Biomed Update 1:100042

9. He T, Mao H, Guo J, Yi Z (2017) Cell tracking using deep neural networks with multi-task learning. Image Vis Comput 60:142–153
10. Kovalevich J, Langford D (2013) Considerations for the use of sh-sy5y neuroblastoma cells in neurobiology. In: Neuronal cell culture. Springer, pp 9–21
11. Lindvall O, Kokaia Z (2006) Stem cells for the treatment of neurological disorders. Nature 441(7097):1094–1096
12. Lu D, Weng Q (2007) A survey of image classification methods and techniques for improving classification performance. Int J Remote Sens 28(5):823–870
13. Nekrasov V, Chen H, Shen C., Reid I (2019) Fast neural architecture search of compact semantic segmentation models via auxiliary cells. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 9126–9135
14. Ounasser N, Rhanoui M, Mikram M, Asri BE (2022) Generative and autoencoder models for large-scale mutivariate unsupervised anomaly detection. In: Networking, intelligent systems and security. Springer, pp 45–58
15. Ounasser N, Rhanoui M, Mikram M, El Asri B (2023) Anomaly detection in orthopedic musculoskeletal radiographs using deep learning. In: Proceedings of eighth international congress on information and communication technology. Springer
16. Ramesh K, Kumar GK, Swapna K, Datta D, Rajest SS (2021) A review of medical image segmentation algorithms. EAI Endorsed Trans Pervasive Health Tech 7(27):e6–e6
17. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 234–241
18. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) Mobilenetv2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520
19. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. ArXiv preprint arXiv:1409.1556
20. Sofroniew MV, Vinters HV (2010) Astrocytes: biology and pathology. Acta Neuropathol 119(1):7–35
21. Yi J, Wu P, Hoeppner DJ, Metaxas D (2018) Pixel-wise neural cell instance segmentation. In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). IEEE, pp 373–377
22. Yi J, Wu P, Huang Q, Qu H, Liu B, Hoeppner DJ, Metaxas DN (2019) Multi-scale cell instance segmentation with keypoint graph based bounding boxes. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 369–377
23. Yi J, Wu P, Jiang M, Huang Q, Hoeppner DJ, Metaxas DN (2019) Attentive neural cell instance segmentation. Med Image Anal 55:228–240

Construction Scheme of Innovative European Urban Digital Public Health Security System Based on Fuzzy Logic, Spectrum Analysis, and Cloud Computing

Yiyang Luo, Vladislav Lutsenko, Sergey Shulga, Sergei Levchenko, and Irina Lutsenko

Abstract In this paper, the growing European data center and the increasingly mature cloud computing technology were proposed to realize the construction of Urban Digital Public Health Security System (UDPHSS). And the scheme of Innovative UDPHSS for European modern cities under the COVID-19 pandemic was designed. Respiratory sounds, containing the structure information of individuals' respiratory system, are analyzed and compared by fast Fourier transform (FFT) and spectrogram. Further extraction and understanding of respiratory sound features of the COVID-19 patients will be aided by means of artificial intelligence. The Pearson correlation coefficients were used to classify individuals' degree of infection with COVID-19, and the membership functions were further constructed to realize the fuzzy logic control of UDPHSS for the allocation of items/medical resources/assistance. And for some famous modern European cities, part of the technical requirements for building such innovative UDPHSS was calculated.

Keywords Artificial intelligence (AI) · Cloud computing · COVID-19 · Fuzzy logic · Pearson correlation coefficient · Portable digital stethoscope · Respiratory sounds · Semi-Markov process · Spectrogram · Fast Fourier transform (FFT) · Urban Digital Public Health Security System (UDPHSS)

Y. Luo (✉) · S. Shulga
V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv 61077, Ukraine
e-mail: yiyangluo@163.com

V. Lutsenko · I. Lutsenko
O.Ya. Usikov Institute for Radiophysics and Electronics of the National Academy of Sciences of Ukraine, 12 Academician Proskura St., Kharkiv 61085, Ukraine

S. Levchenko
International Institute of Applied Research and Technology, Steinbruchweg 2/1, 71069 Sindelfingen, Germany

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_26

1 Introduction

Smartphones were pioneered in the 1990s when IBM engineer Frank Canova realized that chips and chip-and-wireless technology could be put into handheld devices. Along with this, smartphone comes a series of new concepts, including smart city. Despite nearly 30 years of development, digitization has not been completely popularized in modern cities, and our concern for data privacy and protection has always been a limiting factor in the development of urban digitalization, i.e., people have not fully exploited it.

Since 2019, the COVID-19 pandemic has been breaking out all over the world, causing huge damage and impact on people's lives, the economy of society, and the development of countries. As a major public health emergency, the COVID-19 pandemic has six characteristics: (I) the diversity of virus sources, (II) the differences in the spatial–temporal distribution of virus transmission, (III) the extensiveness of virus transmission, (IV) the complexity of causing harm, (V) the complexity of governance, and (VI) the emerge of social-trust crisis. However, this classification does not apply to objective scientific analysis of the impact of COVID-19. The impact of COVID-19 on a city is closely tied to its transmission status within the city.

The spread of the COVID-19 pandemic has prompted municipal leaders to rethink the irreplaceable practical value of ICT and IoT network in improving city management, resilience, sustainability, and more. The construction of the Urban Digital Public Health Security System (UDPHSS) is now a priority and being accelerated. UDPHSS can be used as a social management and control aid during the spreading phase of a pandemic, which helps relieve the pressure of scarce medical resources. Additionally, the implementation of contactless, semi-automated, digital work models in the healthcare sector can provide comprehensive protection and enhance convenience for healthcare workers and volunteers.

2 Concepts and Equipment Basis

2.1 Portable Digital Stethoscope

Existing digital filters mostly use piezoelectric sensors (also known as contact sound sensors) to detect and record respiratory sounds. And piezoelectric sensors, which can directly convert respiratory sound signals into electrical signals through sensitive components, are able to detect lung sounds more effectively than traditional capacitive sensors that rely on air perturbation. In our past work [1–3], the breathing process was compared to a semi-Markov process, the statistical characteristics and spectrograms of the respiratory sounds were studied, and a portable digital stethoscope which can transmit data via Bluetooth was developed, as shown in Fig. 1. The portable digital stethoscope is low-cost and can be linked to a mobile phone via

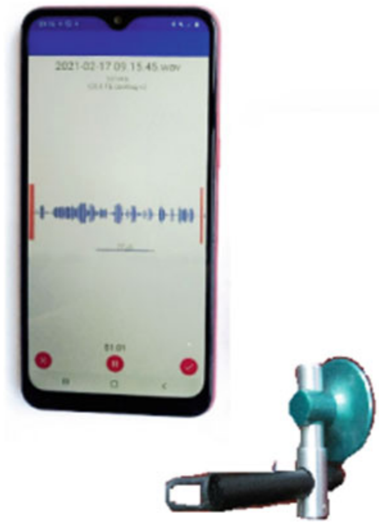


Fig. 1 Portable digital stethoscope designed by Lutsenko team

Bluetooth, and the corresponding mobile phone application has also been designed in a beta version, which is extremely generalizable.

2.2 Auscultation and Quantitative Analysis of Respiratory Sound

The invention of the digital stethoscope has transformed the research on respiratory sounds from a qualitative to a quantitative approach, representing a leap forward in medical technology, making respiratory sounds into data that can be studied and analyzed by data scientists, not just professional doctors. The relationship between breath sounds and the respiratory system is gradually revealed, and the structural information of the respiratory system which contains is fully analyzed. With the help of the recording of respiratory sounds, a non-contact and non-destructive consultation kicked off [1].

From the point of view of system analysis, the respiratory sound as the research object becomes the input of the system, and the whole respiratory system is regarded as a complete nonlinear system containing a variety of explicit variables (age, gender, medical history, lung capacity, etc.) and invisible variables (constriction of the airway, edema, disease condition, presence of foreign bodies, etc.). The respiratory process is an invisible nested semi-Markov process, and the respiratory system that implements this function can be studied and analyzed with the help of structural equation modeling (SEM) [4].

In the process of research, by means of quantitative analysis of respiratory sounds (Pearson correlation coefficient between patients and standards), the condition of patients can be reasonably classified (asymptomatic infection, mild, severe), which constitutes the basic framework of the SEM research system. In this way, the virtual doctor's online consultation function is also realized.

2.3 Detection and Recording of Respiratory Sound

The collection sites of respiratory sound include lung apex, hilum, neck, anterior chest, lateral chest, back, etc. A single-channel wireless Bluetooth digital stethoscope with air-coupled electret condenser microphone can be used to detect and collect breath sounds [2]. Studies on some typical respiratory sounds have shown that the frequency of respiratory sounds is mainly at 100–5000 Hz higher than that of heart sounds at 20–800 Hz. The complex mechanism leads to the complexity of the composition and type of respiratory sounds. The collection of some typical respiratory sounds and the analysis of test data are shown in Table 1 [5]. For our research, the time accuracy and frequency accuracy of the signal are, respectively, required to be $\Delta t \leq 5$ ms and $\Delta f \leq 10$ Hz, which will be regarded as the basic technical requirement indicator used to evaluate the cost and technical requirements of building UDPHSS.

Table 1 Some typical respiratory sounds and the description of their features

Types of respiratory sounds	Main features	Specific feature description	
		Eigenfrequency/ typical frequency	Frequency of energy drop/characteristic duration
Tracheal sound	White noise	100–5000 Hz	800 Hz
Normal (vesicular) lung sound	Low-pass-filtered noise	100–1000 Hz	200 Hz
Bronchial breathing	Strong expiratory component	An intermediate sound, which features between tracheal and normal breathing	
Stridor	Sinusoid	> 500 Hz	
Wheeze		> 100–5000 Hz	> 80 ms
Rhonchus		About 150 Hz	
Squawk		200–300 Hz	About 200 ms
		Followed or preceded by crackles	
Fine crackle	Rapidly dampened wave deflection	About 650 Hz	About 5 ms
Coarse crackles		About 350 Hz	About 15 ms
Pleural friction rub	Rhythmic succession of short sounds	< 350 Hz	> 15 ms

2.4 Fuzzy Logic Control

Fuzzy logic control, referred to as fuzzy control, is a computer digital control technology based on fuzzy set theory, fuzzy linguistic variables, and fuzzy logic reasoning [6, 7]. It has always played an important role in the Internet of things (IoT) and the system control. Diagnosis and classification of medical conditions can constitute fuzzy sets, exemplified by the pandemic COVID-19: {negative, asymptomatic infection, mild infection, severe infection}. For practical application, patients can be classified based on the degree of similarity between their condition and that of typical severe diseases. Then, the corresponding membership function is established to realize the fuzzy logic control of semi-automatic medical resource allocation.

2.5 The Time Complexity

The time complexity is calculated to estimate the technical requirements needed to build an information center of **UDPHSS** for a modern European city. The population, traffic, economic, and medical resource levels of the city itself are also considered as indicators for the calculation.

The time complexity represents the number of operations (expressed as a function) that the algorithm needs to perform and denote it as $f(n)$, where n is data length. The time complexity of the fast Fourier transform (FFT) depends on the data length N and the dimensions M, and for an one-dimensional FFT, the time complexity is

$$f(n) = \mathrm{O}(N * \log N), \tag{1}$$

For an $M * N$ two-dimensional data, the time complexity of FFT is

$$f(n) = O(M * N * \log(M * N)), \tag{2}$$

In past papers [1, 2], it has been found that when the sampling frequency is 44,100 Hz, and the sample length is $2^{16} = 65{,}536$, 94% overlap of realization can be achieved, which is approximately 95% confidence interval, and the reliability is high enough. Taking into account the actual/clinical application, the duration time of each collection was set to 30 s (the sample length is $1{,}323{,}000 \gg 65{,}536$). In order to ensure that the time resolution of the spectrogram does not exceed 10 ms, the signal was divided into 3000 segments, and the length of each segment is 10 ms (441 samples and frequency resolution are 100 Hz). Then, for our research, the time of calculating a Fourier transform is

$$f_{\mathrm{FFT}}(n) = 3000 * 441 * \log(3000 * 441) * O(1) = 1.8648 \times 10^7 * O(1), \tag{3}$$

where the $O(1)$ is a constant order, representing the minimum running time for the computer to perform an operation.

Similarly, the time complexity of choosing the operation and computing the covariance matrix and multiplication is, respectively:

$$f_{\text{ch}}(n) = O(N),\ f_{\text{cov}}(n) = O\left(N^2\right),\ f_{\text{multi}}(n) = O\left(N^2\right). \tag{4}$$

3 Preliminary Scheme of Innovative European Urban Digital Public Health Security System (UDPHSS)

Preliminary scheme of UDPHSS is shown in Fig. 2. The whole scheme consists of four modules:

- **Classifier module**: By using COVID-19 rapid kit detection and the Pearson correlation coefficient (r_i) of spectral characteristics in respiratory sound signals between individuals and typical severe illnesses, users can be classified into four groups: {negative (test line invisible), asymptomatic infection (test line visible, $-1 < = r_i < = 0$, completely inconsistent with typical severe pathological features), mild infection (test line visible, $0 < r_i < \text{mean}(r_i)$, not very consistent with typical severe pathological features), severe infection (test line visible, $\text{mean}(r_i) < r_i < = 1$, consistent with typical severe pathological features)}. The quality of the classification results will directly determine whether the follow-up medical resources can be properly allocated, and the threshold $\text{mean}(r_i)$ needs to be carefully selected;
- **Fuzzy logic control module for medical supplies/aid distribution**: Based on the multiplicative aggregation, the comprehensive score of the user's health status is obtained, and then, a membership function is established to realize the fuzzy logic control of the distribution of medical supplies/aids, and appropriate medical supplies/resources will be given to those who indeed need them (for example: preventive supplies for negative, vitamin tablets for asymptomatic infection, over-the-counter medications for mild infection, first aid and hospitalization for severe infection). The initial membership function is established based on a set of initially collected sample information and is regularly updated to accommodate an increasing and changing sample set;
- **Information processing and storage module of the assistance center**: Spectral characteristics of respiratory sounds of typical severe infection patients, basic information of all users, as well as respiratory sound samples collected recently of infected persons will be stored in the database. This module is responsible for extracting the characteristics of typical severe infections and performing the classification operation. Based on the results, final recommendations, directives, and health certificates will be generated and fed back to both the user and the city management center;

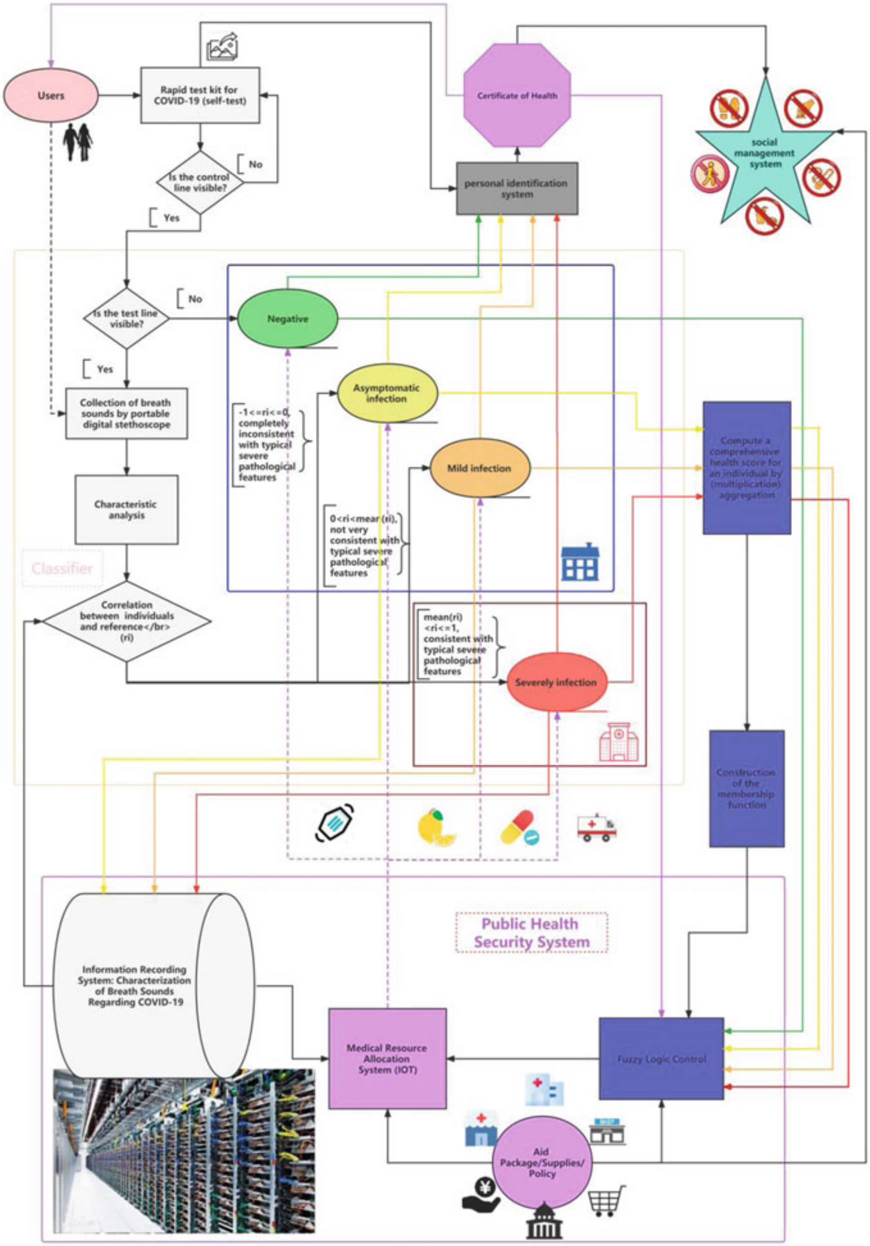


Fig. 2 Preliminary scheme of innovative European Urban Digital Public Health Security System (UDPHSS) based on fuzzy logic, spectrum analysis, and cloud computing

- **Social management module**: After receiving the advice from the information center and the health certificate about the user, the deployment of materials and urban epidemic prevention measures will be implemented. Respond and help with specific individual circumstances and requests. This module can be seen as an output module. In practical applications, UDPHSS will serve as an auxiliary system rather than a mandatory control system to achieve its social and public safety and health protection functions, reducing the burden on the medical system and volunteers and improving the resilience of the city system.

It should be noted that the patient's basic attribute information needs to be considered. For example, the individual's respiratory sounds are age-specific, and according to the statistics [8, 9], the typical breathing rate of a healthy adult at rest is 12–16 breaths per minute. But the individual's physical and psychological conditions, as well as other underlying diseases and pathological factors, can lead to increased breathing and varying noise levels during inhalation.

4 Calculation of the Technical Requirements for Building Such Innovative UDPHSS for Some Famous Modern European Cities

4.1 Choice of Parameters for Building UDPHSS

The premise of UDPHSS is the construction of information/data centers. The digitalization of cities, the popularization of smart phones, and the Internet will be the preconditions for the normal application of UDPHSS. As an online human–computer interaction application, it should adhere to ergonomic principles, such as ensuring fast feedback by providing powerful computing speed to prevent the user's waiting time for results from exceeding 10 s [10, 11].

Considering that the original intention of UDPHSS is to deal with pandemics, such as COVID-19, indicators such as daily new infections in the city/oblast/province and the number of sick people in the city/oblast/province should also be taken into account. A more detailed distribution of medical resources within the city, where possible, should also be fully considered.

Basic technical indicators, such as the site of auscultation, the length of signal collection, the sampling frequency, and the spatio-temporal resolution of the spectrogram, will be utilized. By comparing the Pearson correlation coefficient between the spectrogram characteristics of patients and those of typical severe diseases, the patient's type/group can be determined. This allows for patient classification and implementation of fuzzy logic control for medical resource allocation.

4.2 Assessment of Technical Requirements for Some Modern European Cities

The maximum daily increase in the number of infected people in each city needs to be considered. The number of daily judgments required by the information processing center and the storage requirements of the database need to be comprehensively considered. Based on Formula (1–4), the time complexity $f_{\text{total}}(\text{n})$ of the entire algorithm calculation is

$$f_{\text{total}}(n) = f_{\text{FFT}}(n) + f_{\text{ch}}(n) + f_{\text{cov}}(n) + f_{\text{multi}}(n)$$

$$= \max\left\{O(N * \log N), O(N), O\left(N^2\right), O\left(N^2\right)\right\} = O\left(N^2\right) \tag{5}$$

Recorded data is double type (8-bytes, 64-bit). Each infected person collects respiratory sounds at least twice a day (morning and afternoon), and each time is not less than 30 s. According to the current research, it generally takes at least 8 consecutive days (until recovery). The amount of calculation required will be based on the national population reference urban area and population density. Supermarket supplies are also considered. The specific calculation and statistical results of some modern European cities are shown in Table 2.

Table 2 Specific calculation and statistical results of some modern European cities

No.	Country	City name	Population estimate 2022, person [12]	Built-up land area, square kilometers	Urban population density, per square kilometer	Daily maximum number of new infections of country, person [13, 14]
1	Ukraine	Kharkov	1,485,000	152	9770	38,257
2	Spain	Santa Cruz de Tenerife	506,000	42	12,048	157,034
3	Germany	Stuttgart	1,374,000	184	7467	294,468
4	Bulgaria	Sofia	947,000	80	11,838	9916
5	France	Paris	11,060,000	1102	10,040	428,008
6	United Kingdom	London	11,262,000	671	16,784	192,959
7	Belgium	Brussels	2,203,000	336	6557	11,181
8	Turkey	Istanbul	16,079,000	568	28,308	108,563
9	Italy	Milan	5,488,000	859	2225	220,519
10	Poland	Warsaw	1,963,000	211	546	51,690

5 Conclusions

The envisaged scheme of UDPHSS is in line with actual needs, which can help to increase the resilience of cities and the level of trust in the government. UDPHSC is a bold attempt at using fuzzy logic control in the Internet of medical things (IoMT) field. It has proven to have great development potential and can provide reference suggestions for modern European cities dealing with the COVID-19 pandemic. This includes achieving more efficient use, protection, and conservation of medical resources, as well as laying the foundation for the realistic construction of smart cities.

The development and popularization of data centers laid the foundation for the establishment of UDPHSS in modern European cities and made UDPHSS within reach as part of a smart city.

The future work is to further design the UDPHSS in more detail and calculate the technical and economic parameters required for its construction. To ensure compatibility with other existing medical systems and enable remote surgery and disease monitoring and assistance, the UDPHSS will be designed with the aid of IoMT, information and communication technology (ICT), and 5G/6G technologies. With the help of the global Internet, the combination of GIS and UDPHSS will play a crucial role in the containment of the global pandemic, such as COVID-19.

Acknowledgements This work is supported by the National Academy of Sciences of Ukraine.

References

1. Lusenko V, Lusenko I, Luo Y, Babakov M, Nguyen A (2020) Signature extraction technologies from acoustic noise of the breathing process in lung pathologies. In: 2020 IEEE Ukrainian microwave week (UkrMW), pp 590–593
2. Luo Y, Lutsenko V, Shulgar S, Lutsenko I, Nguyen A (2022) Simulation model of respiratory sound and technology for separating characteristics of pulmonary disease. In: Proceedings of seventh international congress on information and communication technology, ICICT 2022, London, vol 2. Lecture Notes in Networks and Systems, vol 448. Springer, Singapore
3. Lutsenko V, Lutsenko I, Babakov M, Luo Y, Sobolyak A (2019) The use of semi-markov nested processes for the description of non-stationary acoustic noise. Telecommun Radio Eng 78(11):1015–1025
4. Bohadana A, Izbicki G, Kraman SS (2014) Fundamentals of lung auscultation. N Engl J Med 370 (8)
5. Duncan OD (1975) Introduction to structural equation models. Academic Press, New York
6. Leekwijck WV, Kerre EE (1999) Defuzzification: criteria and classification. Fuzzy Sets Syst 108(2):159–178
7. Xu Z, Da Q (2003) An approach to improving consistency of fuzzy preference matrix. Fuzzy Optim Decis Making 2:3–12
8. Nielsen J (1993) Response times: the three important limits. Excerpt from Chapter 5 of Usability Engineering by Jakob Nielsen, Academic Press, AFIPS Fall Joint computer conference, vol 33, 267–277

9. Myers BA (1985) The importance of percent-done progress indicators for computer-human interfaces. In: Proceedings of ACM CHI'85 conference (San Francisco, CA, 14–18 April), 11–17
10. Rodríguez-Molinero A, Narvaiza L, Ruiz J, Gálvez-Barrón C (2013) Normal respiratory rate and peripheral blood oxygen saturation in the elderly population. J Am Geriatr Soc 61(12):2238–2240
11. Gavriely N, Palti Y, Alroy G (1981) Spectral characteristics of normal breath sounds. J Appl Physiol Respir Environ Exerc Physiol 50(2):307–314
12. Demographia World Urban Areas 18th annual 2022.07 http://www.demographia.com/d-new.htm
13. REUTERS COVID-19 Tracker. https://graphics.reuters.com/
14. JHU CSSE COVID-19 Data. https://github.com/CSSEGISandData/COVID-19

Virtual Training System for a MIMO Level Control System Focused on the Teaching-Learning Process

Santiago Zurita-Armijos, Andrea Gallardo, and Victor H. Andaluz

Abstract The present project is centered on the development of a Virtual Environment to perform the application of control algorithms at a MIMO level control learning module. The mathematical model is obtained heuristically and to get the dynamic variables of the system an identification and optimization algorithm is implemented. The environment has been developed using CAD tools with UNITY 3D graphic engine, and MATLAB has been used for the control schemes. The purpose of the developed system is to be a functional tool focused on the teaching-learning process in the industrial automation field that is safe and low-cost, especially when access to physical equipment is limited or does not exist.

Keywords Virtual environment · Learning module · Heuristic model · Process simulate

1 Introduction

Automation is a process of economic, social, cultural and technological transformations applied to industry, where the production of goods was carried out in a mechanized manner [1], currently represents a great advantage by increasing productivity and reducing labor costs [2]. The industrial revolution played its significant role since the era of industrialization began in the 18Th century with mechanical equipment driven by the power of water and steam [3]; the second industrial revolution occurred in 1870, when electric power formed a major system known as mass production [4]; during the 1970s, the third industrial revolution occurred with the

S. Zurita-Armijos (✉) · A. Gallardo · V. H. Andaluz
Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador
e-mail: sdzurita@espe.edu.ec
URL: https://www.espe.edu.ec

A. Gallardo
e-mail: aegallardo1@espe.edu.ec

V. H. Andaluz
e-mail: vhandaluz1@espe.edu.ec

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_27

rise of electronics [5]. Business development linked to industrial intelligence, robots and virtual environments has driven the emergence of Industry 4.0, referred to as the fourth industrial revolution, which describes the merger of the factory with the Internet through the design and implementation of intelligent components with the virtualization of systems and processes to achieve greater flexibility and individualization of production processes [6].

As part of Industry 4.0, Virtual Reality (VR), can conceive any type of environment, redesign it, test it and refine it in a virtual computing scenario [7]. Is the science of how to convert a physical object or resource to an emulated or simulated object in software [8], with technologies capable of combine elements from the real world with elements from the virtual world [9]. Virtual training systems are aimed at solving problems in a practical way [10], in which training new personnel is linked to constraints such as high-cost training programs, difficulties in accessing the physical environment, and human error in the development of their activities, among other things. They are a great tool for learning new skills in different areas, such as in [11–13] where robotic assemblies are virtualized in industrial scenarios; training systems for critical procedures that require a lot of knowledge on the part of the operator, such as in [14, 15]; in the area of automation education, [16] develops a virtual system to simulate control algorithms in an assistive chair for persons with reduced mobility.

The aim of virtualizing a process is to provide a sense of immersion and interaction to capture the user's attention. Ruiz et al. [17], useful features that can be used for training or teaching workers or students. 2020 was a different year, due to the pandemic where isolation and disease brought the world to a virtual standstill [12], the classrooms were emptied, the lockdowns were instated, the educational system quickly changed to emergency distance learning [18], which caused both students and teachers to adapt to new technologies that would allow them to achieve their academic goals. Professionals in the field of education recognize that the teaching-learning process requires the use of information and communication technologies (ICTs) [19], which is why we must currently prioritize the development of systems that provide teaching tools with appropriate pedagogical criteria for the training of students. As previously indicated, this work proposes the development of a virtual training system of a level control plant, designed, built and mathematically modeled for the application of advanced control algorithms to be used in the process of teaching advanced control to engineering students.

2 Formulation of the Problem

The virtualized systems in the industry allows users to interact with the processes to learn how they work, devise possible improvements, plan maintenance, and instruct the operator in a safe and realistic environment without physical interaction with the processes. Training and educating workers are two of the best strategies to control

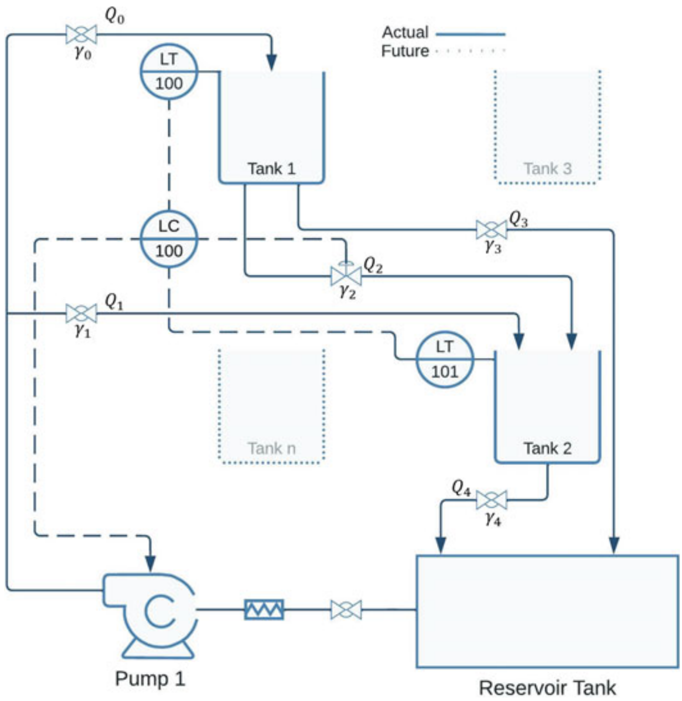


Fig. 1 MIMO level control process P&ID

and reduce the accident rate [9], a very critical factor in the industry. Within a massive industry-wide change VR becomes a fundamental pillar because it shows promise in the area of error diagnostics and training [20].

The importance of a quality e-learning plan has recently become evident, especially in areas where it is difficult to replace in-person education with the virtual one [21], mainly when laboratories or interacting with equipment are required. Is in this area where the virtualization of models becomes a very important tool because visually students can interact with the elements and know how they work. Therefore, we propose the design of a virtual environment focused on the application of MIMO level control algorithms for the teaching-learning process in the field of industrial automation, the P&ID is represented in Fig. 1.

For the formulation of the mathematical model describing the dynamic behavior of the process, it is taken into account that the volume of the liquid in the plant undergoes variation in time due to the inflow and outflow, as expressed in (1).

$$\frac{\mathrm{d}V(t)}{\mathrm{d}t} = Q_{\text{in}} - Q_{\text{out}} = A_i \frac{\mathrm{d}h_i(t)}{\mathrm{d}t} \tag{1}$$

where V represents the tank volume; Q_{in} and Q_{out} represent the inlet and outlet flow rates respectively; A_i when $i = 1, 2$ represent the cross-sectional area of the tanks; h_i when $i = 1, 2$ represent the height of the tanks. Tanks 1 (T1) and 2 (T2) are considered to have a constant cross-sectional area and the height of the liquid inside them is considered to vary over time. Based on (1) it is possible to determine the behavior of the level of T1 according to the mass balance expressions represented in Fig. 1.

$$A_1 \frac{\mathrm{d}h_1(t)}{\mathrm{d}t} = Q_0 - Q_2 - Q_3 \tag{2}$$

$$\frac{\mathrm{d}h_1(t)}{\mathrm{d}t} = \frac{\gamma_0 \left(1 - \gamma_1\right) k_1 v(t) - \left(k_2 \gamma_2(t) + k_3 \gamma_3\right) \sqrt{2\, g h_1(t)}}{A_1} \tag{3}$$

where Q_0 represents the inlet flow rate; Q_2 and Q_3 represent the outlet flow rates of T1; γ_i when $i = 0, 1, 2, 3, 4$ represent the setting values of the valves (0 is completely closed, 1 is completely open); k_i when $i = 1, 2, 3, 4$ represent the gains of the pump and valves respectively; $v(t)$ represents the voltage applied to the pump; g represents the gravity. Similarly for T2, Q_1 and Q_2 represent the inlet flow rates and Q_4 the outlet flow rate. The behavior of the T2 level is represented by:

$$A_2 \frac{\mathrm{d}h_2(t)}{\mathrm{d}t} = Q_1 + Q_2 - Q_4 \tag{4}$$

$$\frac{\mathrm{d}h_2(t)}{\mathrm{d}t} = \frac{\gamma_1 k_1 v(t) + k_2 \gamma_2(t) \sqrt{2\, g h_1(t)} - k_4 \gamma_4 \sqrt{2\, g h_2(t)}}{A_2} \tag{5}$$

3 Methodology

The methodology used for the development of the system is shown in Fig. 2, showing the stages for the implementation and development of the virtual MIMO level control environment. The proposed system has three main stages: (i) Conceptualization, in this stage, the design is carried out, which consists of interconnected tanks, loading valves, a motorized valve and a pump to fill the tanks. With the arrangement of the elements, mathematical modeling simulating the dynamic behavior of the process is obtained as shown in the expressions (3 and 5) it is necessary to consider the inflow and outflow rates of each of the T1 and T2 tanks. Is important to note that, by changing the opening value of the valve γ_0, process operation can be made independent in case SISO control algorithms need to be tested. Additionally, identification and validation algorithms are implemented to obtain the dynamic variables of the system, this process is performed through experimental tests. (ii) Scheme Controls, the system allows the application of control algorithms focused on the learning process, in this paper is intended to control the level of the interconnected tanks T1 and T2 via control signals for the activation of the pump and the motorized valve. (iii) Virtualization, the physical elements of the module are designed with CAD tools;

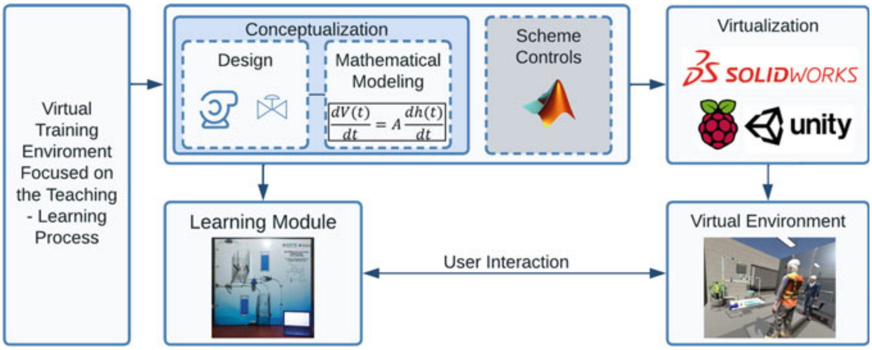


Fig. 2 Methodology for implementation of control algorithms using virtual environments

additional functionality of the load valves is included to add disturbances to the system; after the design, Solidworks Visualize software is used (Dassault Systèmes SolidWorks Corporation) to render the model, add textures, lighting, color and export it to a file supported by Unity 3D software; the communication between the PC with the controller and the virtual environment is done through MODBUS TCP/IP. Finally the tests are developed between the learning module and the virtual environment, which allows to check and compare the operation of the control algorithms in the physical system as well as in the virtual one; in order to validate the operation of the virtual environment.

3.1 Virtualization

The virtual environment developed is centered on the learning process, it must contain elements that allow the learner to interact with, realistic scenarios of an engineering laboratory, and sounds that increase the level of realism to motivate the student to perform their tasks of implementation of control algorithms. The procedure of the realization of the virtualized learning module is shown in Fig. 3.

Four main stages are established for the design of the virtual environment: (i) External Resources, here are included all the related elements that will be found within the virtual platform, these items can be arranged in two groups: (a) Learning Module, within this are located the elements with which you can interact, in this case, the tanks, valves and pump, it is here where you can visually identify the control algorithms that will be implemented. (b) Users and Environments, it is here where the space where the student and the teacher will navigate is designed, it is set in a laboratory where industrial automation practices are performed within an educational institute. The design of the virtual environment is carried out in a CAD tool with many capabilities such as 3D Solidworks [22], the geometries of the system are modeled and the movements of the animations are restricted. Solidworks Visualize is used

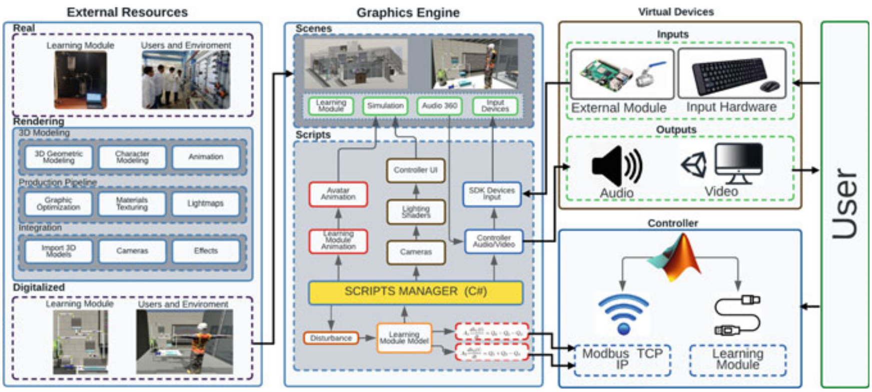


Fig. 3 Proposed outline of the virtual environment

for rendering, adding textures, materials, graphic optimization and for exporting to a Unity 3D compatible file. (ii) Graphics Engine, here is developed all the emulation of the real learning module process, it can be divided into two groups: (a) Scenes, which consists of everything designed in stage (i), audio of pump ignition, valve opening, environmental noise and all the elements that allow the user to participate in an interactive experience, so typical elements to be found in a real learning laboratory are added, simulations of the visual change of the level of the tanks are included with which the performance of the control algorithms to be applied can be reviewed; a real-time trend graph is included in which the heights of T1 and T2 and the control actions of the motorized valve and the pump are monitored. (b) Scripts, these are developed in C Sharp (C#) and must be added to a scene object so that they can be called by Unity [23], these can be used to emulate the system compartment, movements and the change of tank level variables; additionally, there is a script that allows interaction with an external module that was built to allow valve opening and closing, this action is represented visually within the virtual environment, they are used to add disturbances to the mathematical model of the process, in addition to increasing the user intervention with the environment. The remaining scripts manage the interface, cameras, alarms and sounds of the elements of the virtual learning module. (iii) Controllers, are the algorithms that are applied in the real and virtual learning module, the desired values are the heights of T1 and T2 modifying the control actions that directly affect the motorized valve and the pump, within Matlab will be the design of the controllers, they will communicate via MODBUS TCP IP to a Raspberry Pi 4 board containing the mathematical model of the system in Phyton, this, in turn, sends the variables of the variables to be affected in the simulation to Unity. For the physical learning module, Matlab sends the control actions via serial communication to an Arduino Uno R3 board. (iv) Users, who are in charge of interacting directly with the modules, implementing the control algorithms, selecting the Set Point (SP) values, adding disturbances to the system and verifying its operation.

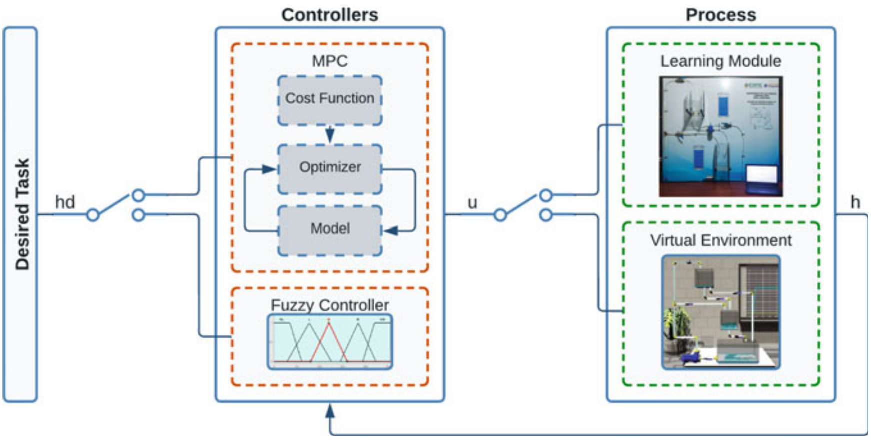


Fig. 4 Suggested controller scheme

4 Control Schemes

This section describes the formulation of the controllers; Fuzzy and Model Predictive Control (MPC); its objective is to maintain the tank head at desired values (SP) by manipulating the voltage level applied to the pump and the motorized valve. The control schemes are subject to disturbances that can be added with manual loading valves as indicated in Fig. 4.

4.1 Fuzzy Controller

The fuzzy logic-based level control system has two inputs, the voltage applied to the Pump and the opening of the Motorized Valve; and two outputs, height T1 and height T2, for the implementation of Outputs Membership Functions (OMF) the operating range is divided into seven, with trapezoids at the ends and triangles in the center, operative ranges for both OMFs are [0–30] centimeters, which is the actual measurement of the tanks of the learning module; each is constituted of the following sets: VVL (very very low), VL(very low), L(low), M (medium), H(high), VH(very high), VVH (very very high). For Inputs Membership Function (IMF) in a similar way both, the pump and the valve have five operating ranges, with trapezoids at the ends and triangles in the center, operative range of the pump output is [0–255] PWM and it is constituted of the following sets: VS(very slow), S(slow), M(medium), F(fast), VF(very fast); and for the valve its operative range is [0–1] and is constituted of the following sets: VC(very close), C(close), M(medium), O(open), VO(very open). In Fig. 7 the IMFs and OMFs are shown. Twenty-five rules were established which were developed based on the experience gained when testing was performed.

4.2 Model Predictive Control

Within the present paper a nonlinear MPC control scheme is used, where the system can be expressed as:

$$\dot{h}(t) = f(u(t), h(t)) \tag{6}$$

where, $\dot{h(t)} = \begin{bmatrix} \dot{h_1} & \dot{h_2} \end{bmatrix} \in \Re^2$ represents the vector of the rate of variation of the outputs of the process to be controlled; $h(t) = \begin{bmatrix} h_1 & h_2 \end{bmatrix} \in \Re^2$ representing the outputs to be controlled and $u(t) = \begin{bmatrix} \upsilon & \gamma \end{bmatrix} \in \Re^2$ which represents the maneuverability vector of the pump and motorized valve actuators respectively.

When applying the predictive control algorithm, the cost function (J) is defined which depends on the control error and the changes in the control actions, where control error is defined by:

$$e(k + i \mid \mathrm{k}) = h(k + i) - h_d(k + i \mid \mathrm{k}) \tag{7}$$

Over a prediction horizon N and the sum of the norm of the predicted increments of the control action over a control horizon N_u:

$$J = \sum_{i=1}^{N} \delta_i \|e(k + i \mid k)\|_Q^2 + \sum_{i=1}^{N_u} \lambda_i \|\Delta u(k + i - 1 \mid k)\|_R^2 \tag{8}$$

where, K represents the current time instant; i represents the sampling period to be predicted; hd represents the desired values; δ_i and λ_i are sequences chosen as penalty constants [24]. In this way, J can be specified as a function that depends solely on future control actions (Fig. 5).

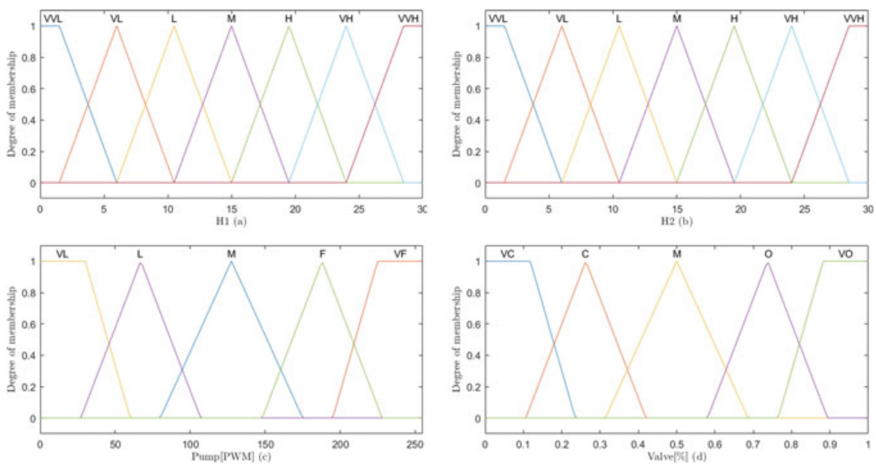


Fig. 5 Membership functions. **a** IMF $H1$, **b** IMF $H2$, **c** OMF Pump and **d** OMF valve

Fig. 6 User interaction with the built learning module

5 Analysis and Results

Experimental outcomes are described in three subsections: (5.1) presents the construction of the real learning module from which the mathematical model was obtained, the user's interaction with the learning module is shown in Fig. 6; (5.2) indicates the details of the virtual environment designed for the learning module and (5.3) presents the implementation of the proposed controllers.

5.1 Learning Module Construction

The Learning Module was developed based on the P&ID of Fig. 1 which consists of two tanks (T1 and T2) interconnected using a motorized valve (γ_2), a reservoir tank that stores the liquid to be driven by a direct current (DC) pump, which is responsible for carrying the fluid to T1 and T2, three manually activated valves that allow adding disturbances to the system and a set of manual valve and filter for maintenance of the reserve tank. The system includes two ultrasonic-type sensors to quantify the liquid level in each of the tanks.

The system structure was designed to be scalable, i.e., to add or remove module elements or redistribute them, so that new configurations can be tested to obtain new mathematical models, or to test new control algorithms according to user requirements and needs. The distribution of the elements of the system is shown in Fig. 7.

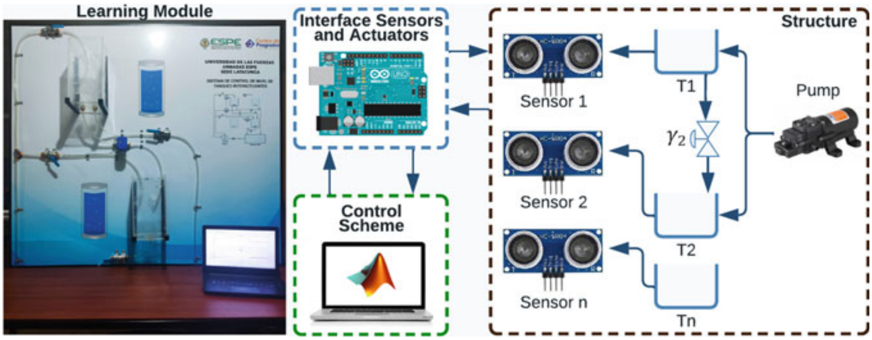


Fig. 7 Learning module design elements

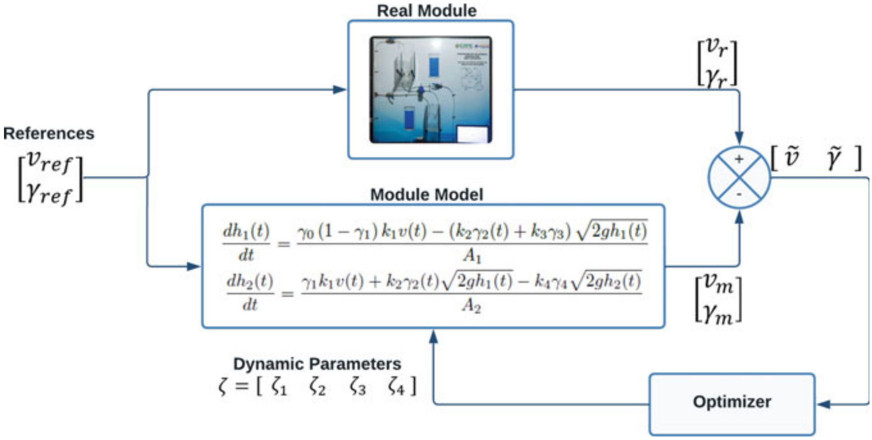


Fig. 8 Identification scheme

For the identification of the dynamic model of the process, experimental tests were performed with the real system, the data recorded were introduced in the identification algorithm, which made it possible to obtain the dynamic variables of the system by means of an algorithm based on optimization and validation by comparison with the mathematical model of the process. The procedure carried out is as follows: (i) Excitation of the module, with different steps of both the pump and the motorized valve to know the output of the system at certain input values at a certain instant of time; (ii) Identification algorithm, the variables of the model are identified from the data collected in the prior step. To reduce the error, an optimization algorithm is implemented that compares the values obtained from the module with those obtained with the mathematical model as indicated in Fig. 8.

Where, $\zeta = \left[\zeta_1 \ \zeta_2 \ \zeta_3 \ \zeta_4 \right]$ represents the vector of pump (k_1) and valve gains (k_2, k_3 y k_4) from Eqs. (3) and (5).

5.2 Virtual Environment

In Fig. 9a the real module and the virtualized module are visualized, it is observed that they have similar proportions and the same amount of control and monitoring elements, they have the same layout to perform the scaling of the system. The virtual environment where the training module is located is designed in such a way that it resembles an engineering laboratory; Fig. 9b, has training modules, rest areas for students, shelves and furniture that resemble a teaching area. The area where the virtual module is located is set in a practice room, at the back of which there is a blackboard showing the trend graphs of the variables inherent to the process; Fig. 9c. Two main user levels are established; Fig. 9d, the student-user allows to visualize the operation and implement the control algorithms; teacher-user can carry out the opening of all the manual valves causing disturbances to the system that will directly influence the operation of the controller. The user-level teacher can manipulate the valves; Fig. 9e, when interacting with these elements within the virtual environment, the closing animation can be seen and a sound is played as a result of the opening or closing action. In the event of an overflow of the liquid in the tank due to an error in the control scheme, the fluid will start to fall from the tanks, alarm sounds are played and the teacher enters the room to check the status of the system, Fig. 9f. The system was developed on a computer with an AMD Ryzen 7 5800H processor, NVIDIA RTX 3060 graphics card and 16.0 GB of RAM.

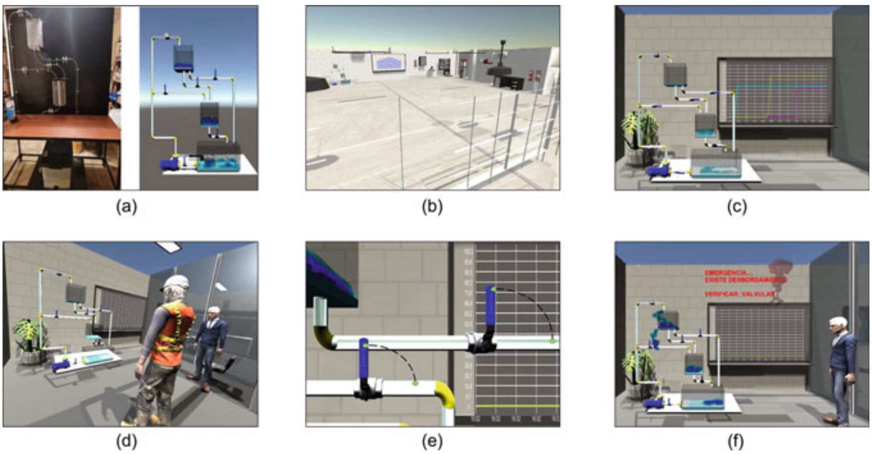


Fig. 9 Learning Module. **a** Real—Virtualized Module Comparison, **b** Laboratory Environment, **c** Module and Trends, **d** Users, **e** Valves Manipulation and **f** Tanks overflows

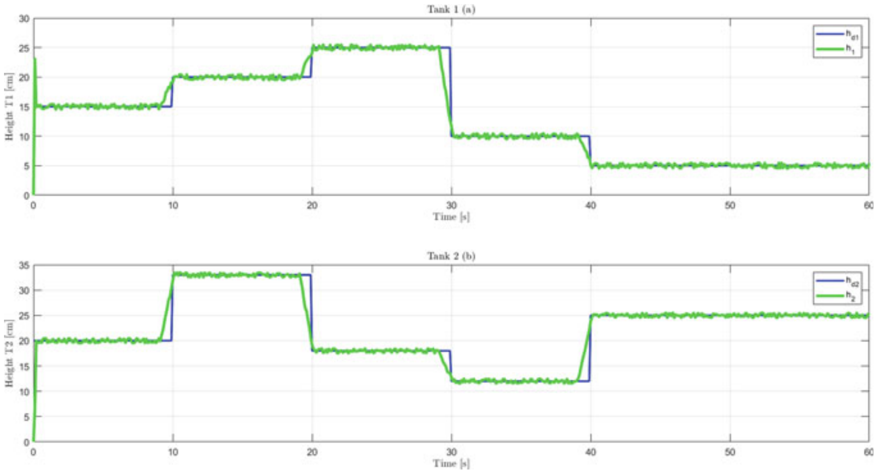


Fig. 10 System evolution in the virtual environment. **a** Tank 1, **b** Tank 2

5.3 Implemented Control Schemes

For the experimentation with the virtual environment, the mathematical model was used with the gains obtained in the identification and optimization process; a Matlab script was used for the simulation process. Figure 10 shows the response of the system at different desirable height levels for tanks T1 and T2.

Disturbances were added to the system output to verify the behavior of future control actions, very small errors were obtained and in all cases, the desired height value was reached in both T1 and T2. Similar results were obtained in both control schemes (Fuzzy and MPC).

6 Conclusions

The development of the virtual training module for a MIMO level control oriented to teaching and learning processes has been successful in simulating the performance of the real learning module that was developed for engineering practices, allowing to copy its behavior and emulate it in an environment that is reliable, safe, and free of risk for people. Provides a functional tool for future projects in the field of industrial automation for the implementation of multi variable control algorithms. The experimentation with the mathematical model has made it possible to test the efficiency of the virtual training module, it was necessary to apply an algorithm of identification and optimization to obtain the dynamic parameters of the system for which tests were made to the physical learning module. The designed controllers allow correcting the disturbances produced by manipulating the load valves present in the modules. The system is scalable, i.e., the arrangement of the elements can be changed or, if required, increased or decreased.

Acknowledgements This article shows the results of the Degree Project of the Master's Program in "Maestría en Electrónica y Automatización con Mención en Redes Industriales" of the Postgraduate Center of the Universidad de las Fuerzas Armadas ESPE.

References

1. Rozo-García F (2020) Revisión de las tecnologías presentes en la industria 4.0.: Revista UIS Ingenierías (in Spanish), 19(2):177–192. https://doi.org/10.18273/revuin.v19n2-2020019
2. Moradiya MA (2018) An: introduction to automation in industry. Azorobotics
3. Xu M, David JM, Kim SH (2018) The fourth industrial revolution.: opportunities and challenges. Int J Financ Res 9(2). https://doi.org/10.5430/ijfr.v9n2p90
4. Kanji GK (1990) Total quality management the second industrial revolution. Total Qual Manage 1(1). https://doi.org/10.1080/09544129000000001
5. Taalbi J (2019) Origins and pathways of innovation in the third industrial revolution. Ind Corp Change 28(5). https://doi.org/10.1093/icc/dty053
6. Peralta-Abarca J, del C, Martínez-Bahena B, Enríquez-Urbano J (2020) Industria 4.0. Inventio 16(39). https://doi.org/10.30973/inventio/2020.16.39/4
7. Liagkou V, Salmas D, Stylios C (2019) Realizing virtual reality learning environment for industry 4.0. Procedia CIRP 79. https://doi.org/10.1016/j.procir.2019.02.025
8. Kihara T, Muli E, Chege L (2019) Adoption of virtualization by government organizations in Kenya. 2019 IST-Africa week conference (IST-Africa), pp 1–10. https://doi.org/10.23919/ISTAFRICA.2019.8764846
9. Zambrano JI, Bermeo DA, Naranjo CA, Andaluz VH (2020) Multi-user virtual system for training of the production and bottling process of soft drinks. In: 2020 15th Iberian conference on information systems and technologies (CISTI). IEEE, pp 1–7
10. Liu Y, Sun Q, Tang Y, Li Y, Jiang W, Wu J (2020) Virtual reality system for industrial training. In: 2020 International conference on virtual reality and visualization (ICVRV), pp 338–339. https://doi.org/10.1109/ICVRV51359.2020.00091
11. Martins A, Costelha H, Neves C (2019) Shop floor virtualization and industry 4.0. In: 2019 IEEE international conference on autonomous robot systems and competitions (ICARSC), pp 1–6 (2019). https://doi.org/10.1109/ICARSC.2019.8733657
12. Beloiu R (2021) Virtualization of robotic operations. In: 2021 12th international symposium on advanced topics in electrical engineering (ATEE), pp 1–4. https://doi.org/10.1109/ATEE52255.2021.9425336
13. Cobo EB, Andaluz VH (2021) Virtual training system for robotic applications in industrial processes. International conference on augmented reality, virtual reality and computer graphics. Springer, Cham, pp 717–734
14. Petkov E, Angelov V (2020) Virtual reality training system for specialists who operate on high-voltage Switchgears in an oil plant in Russia. In: Proceedings of the 21st international conference on computer systems and technologies' 20, pp 266–269. https://doi.org/10.1145/3407982.3408003
15. Romo JE, Tipantasi GR, Andaluz VH, Sanchez JS (2019) Virtual training on pumping stations for drinking water supply systems. In: International conference on augmented reality, virtual reality and computer graphics, Springer, Cham, pp 410–429. https://doi.org/10.1007/978-3-030-25999-0_34

16. Ortiz JS, Palacios-Navarro G, Andaluz VH, Guevara BS (2021) Virtual reality-based framework to simulate control algorithms for robotic assistance and rehabilitation tasks through a standing wheelchair. Sensors 21(15):5083. https://doi.org/10.3390/s21155083
17. Ruiz RJ, Saravia JL, Andaluz VH, Sánchez JS (2022) Virtual training system for unmanned aerial vehicle control teaching-learning processes. Electronics 11(16):2613. https://doi.org/10.3390/electronics11162613
18. Vrgović P, Pekić J, Mirković M, Anderla A, Leković B (2022) Prolonged emergency remote teaching: sustainable e-learning or human capital stuck in online limbo? Sustainability 14(8):4584. https://doi.org/10.3390/su14084584
19. Guillén-Gámez F, Cabero-Almenara J, Llorente-Cejudo C, Palacios-Rodríguez A (2021) Differential analysis of the years of experience of higher education teachers, their digital competence and use of digital resources: comparative research methods. Tech Knowl Learn 1–21. https://doi.org/10.1007/s10758-021-09531-4
20. Almeida LSdO, Lugli AB, Pimenta TC, Silva MVCe, Henriques JPC, Mesquita RP (2020) Virtualization of an aluminum cans production line using virtual reality. In: 2020 27th international conference on mixed design of integrated circuits and system (MIXDES), pp 282–287. https://doi.org/10.23919/MIXDES49814.2020.9156023
21. Hinojosa CJT, Cabrera JJF, Mora HRC, Garzón NVO (2021) An augmented reality based E-learning tool for engineering. In: 2021 IEEE colombian conference on communications and computing (COLCOM), pp 1–6. https://doi.org/10.1109/COLCOM52710.2021.9486294
22. Diseño/Ingeniería (in Spanish): https://www.solidworks.com/es/domain/design-engineering. Last accessed 21 July 2022
23. Coding in C# in Unity for beginners. https://unity.com/es/how-to/learning-c-sharp-unity-beginners. Last accessed 14June 2022
24. Andaluz GM et al (2016) Modeling dynamic of the human-wheelchair system applied to NMPC. In: Kubota N, Kiguchi K, Liu H, Obo T (eds) Intelligent robotics and applications. ICIRA 2016. Lecture Notes in Computer Science, vol 9835. Springer, Cham. https://doi.org/10.1007/978-3-319-43518-3_18

Machine Learning Prediction of Intellectual Property Rights Based on Human Capital Factors

Chasen Jeffries and Karina Kowarsch

Abstract Regression modeling approaches with sufficient literature support have postulated that intellectual property rights (IPR) have a positive impact on human capital in general. However, few papers attempt to uncover the impact of human capital on IPR protections. As IPR fosters innovation, it is critical to understand how developments in education, technology, and health care affect IPR. This paper primarily focuses on the investigation of what specific human capital indicators would be good predictors of IPR and is conducted using machine learning techniques. Compared to ridge regression and pruned regression tree, the random forest model outperforms all other models, with the highest R squared score and the lowest RMSE score. The random forest model suggests that among health, technological skills, and education, university enrollment per capita and physicians per capita play a more important role for predicting intellectual property rights. Moreover, using classification modeling techniques, a neural network model with a few hidden layers and less elements in each hidden layer effectively overcomes the overfitting issue and surpasses all other more complex neural network models. This finding indicates that the concision and precision of artificial intelligence models helps simplify the degree of complexity of social science.

Keywords Human capital · Intellectual property rights · Machine learning · Artificial neural network

C. Jeffries (✉) · K. Kowarsch
Claremont Graduate University, Claremont, CA 91711, USA
e-mail: chasen.jeffries@cgu.com

K. Kowarsch
e-mail: dandan.kowarsch@cgu.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_28

1 Introduction

Intellectual property protections and human capital are critical components of long-run economic growth. IPR provide protection for patents, copyrights, trademarks, and trade secrets that are fundamental to incentivizing R&D. Human capital interacts with R&D to innovate creating new technologies and pushing steady-state output higher. Human capital has two primary components: educational attainment and learning by doing. Educational attainment has a number of measures including secondary school rate and average years of total schooling. Learning by doing often focuses exclusively at on-job training and experiences that increase the human capital of the workers, but it can include learning outside of both school and work (e.g., learning by reading a technical book). Previous regression analysis found IPR to have a positive effect on human capital accumulation [1]. While previous literature has briefly investigated the interactions of human capital and IPR protections [2], this article seeks to investigate the impact of human capital factors on intellectual property rights protections. Does human capital attainment stimulate increased IPR protections? Which forms of human capital in an economy have the greatest effect on IPR protections? This article uses machine learning to investigate the channels of human capital including: university enrollment, physicians per capita, internet access, and personal computers access that affect IPR protections.

2 Previous Literature

The publication of exogenous growth theory established an economic model based on three factors: capital, labor, and savings [3]. This model highlighted the concept of conditional convergence that macroeconomic growth rate, *ceteris paribus*, is based on the distance away from steady-state output. By assuming exogenous technological growth, it left innovation outside the model, a key constraint that the endogenous growth models attempted to overcome. Romer built a model that internalized technological innovation and human capital [4]. By including technology within the model, it emphasized the importance of incentives for human capital accumulation and technological development in pushing steady-state output higher. Researchers have continued to build on Romer's model by highlighting additional avenues by which innovation and human capital affect economic growth [5].

The innovation literature argues that IPR protections provide microeconomic incentives for R&D that is the backbone of technological innovation [6]. R&D expenditures would lack economic feasibility without these protections. Basic economics identifies that R&D is expensive, and without an expected return on investment higher than the expenditure no actor would choose to innovate for economic reasons [7]. IPR protections provide short-term protections that attempt to balance the return on investment with the optimal social gain through technology diffusion [8]. Multiple studies have found IPR protections to have a strong positive effect on economic growth in high- and low-income countries [9], but failed to identify the same effect for middle-income countries [10].

Human capital's role in economic growth has evolved with the shifts in economic model paradigms, but it has always been identified as a positive factor in economic growth. Theory states that human capital is crucial in both acquiring the use of technology and technological innovation [11]. Nations with greater human capital can internalize and implement new technologies at a faster rate [12]. This internalization can simply be imitation of known technologies, a critical aspect of catch-up, but can also include evolutions of or innovations into new technologies. These human capital effects should drive a desire for IPR protections to secure returns on investment for these innovations. Human capital accumulation literature highlights several pathways for its positive effect on IPR protections.

Previous empirical investigations on IPR and human capital have largely ignored the association between these two variables [13]. The research primarily investigates how these concepts effect other economic growth variables, most commonly GDP per capita. These analyses will sometimes include the other variable (IPR or HC) as a control variable. Ginarte and Park [14] developed an index for IPR protections and evaluated the impact of variables in patent protections. Their research failed to identify a positive effect of human capital on patent protections. Loukil [15], as well as Sorck and Diwakar [16], identified that higher human capital nations are better equipped to use IPR protections than lower human capital nations. Chen and Puttitanuns [17] fail to find support for a positive association of human capital on IPR protections when making use of a two-tail test. Gould and Gruben find that human capital has explanatory power toward IPR [18].

Overall, the economic paradigm has postulated several causal pathways for human capital and IPR protections to be associated. Empirical investigations have failed to sufficiently investigate these theories and their theoretical implications, with few analyses including both variables in the same model. The limited research has mixed results about the effect of human capital on IPR protections.

3 Research Design

3.1 Descriptive Analysis

The sample dataset is pulled from CNTS 2020 [19] and Ginarte and Park IPR Index [2]. The sample data are a quasi-panel data frame since the time frame is discrete. In the dataset only certain years are included (1960, 1965, 1970…2005). Considering this characteristic of the sample data, rather than treat the sample data as cross-sectional time series dataset, we handle the sample as cross-sectional data and set country name and year as controlled variables.

The sample data have missing values that we replace with zeros. The final dataset includes 945 observations for each attribute. The dependent variable is IPR score

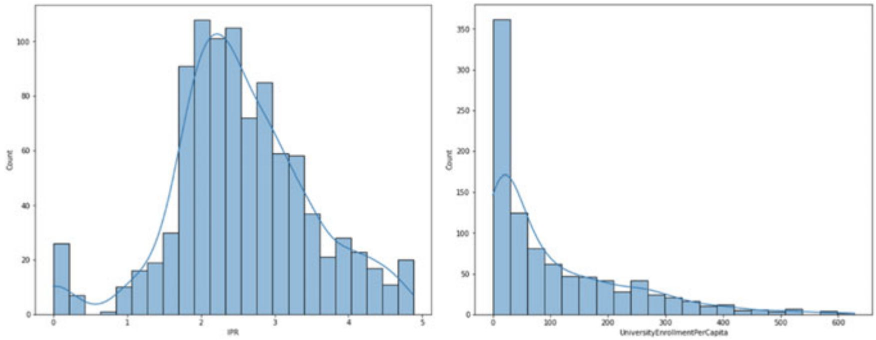


Fig. 1 Univariate analysis on IPR protections (left) and university enrollment per capita (right) using histogram and boxplot

for the regression models or the binary variable of IPR (high and low) for the classification models. The explanatory variables are: university enrollment per capita, physicians per capita, book production by title per capita, internet users per capita, personal computer per capita, and passenger cars per capita. The control variables include but are not limited to gross national income (GNI) per capita and population. The sample consists of 103 countries. The top frequency of country counts is 10, while the lowest frequency is 4.

Univariate and Bivariate Analysis on Numerical Data Type

We begin by examining the distribution of IPR and university enrollment. Figure 1 displays the histogram combined with boxplot of the attributes IPR and university enrollment per capita in this sample dataset. The histogram of IPR (left) suggests that the variable is close to normal, but outliers are identified for this variable. We find the shape of university enrollment per capita (right) is highly right skewed, meaning most data points are clustered at low end, with only a few data points on the far right.

We are also interested in the attribute of physicians per capita that is identified as an indicator of a high skill workforce. We examine one control variable, GNI per capita, that is, expected to have a strong impact on IPR. Figure 2 presents the distributions of physicians (left) and GNI (right).

The histograms show that both variables have upper outliers and are highly right skewed. Clearly, the distributions of both attributes are highly alike. Even though we uncovered the outliers in the attributes, we keep them unchanged. We will implement an advanced machine learning model to resolve the challenges that may be caused by outliers.

Next, we check the correlation between the variables of interest. The scatterplot of IPR and university enrollment and IPR and physicians are displayed in Fig. 3. Both university enrollment and physicians per capita appear to be positively correlated with IPR.

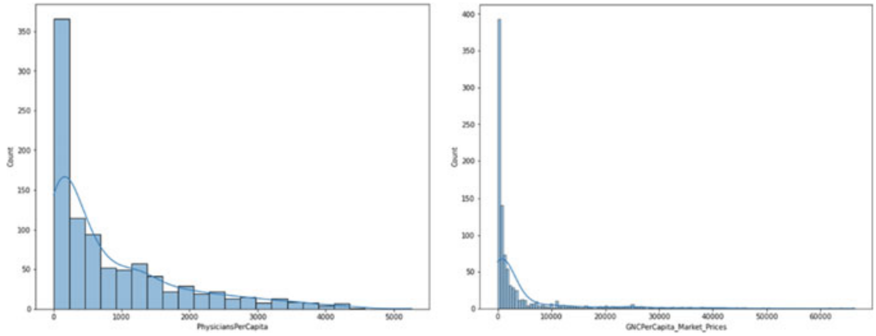


Fig. 2 Distribution of physicians per capita (left) and GNI (right) per capita using histogram and boxplot

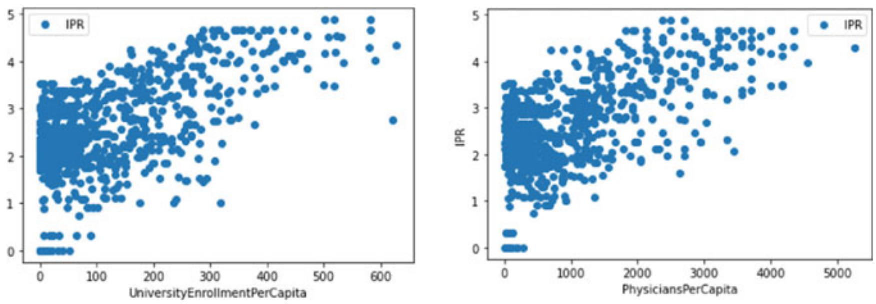


Fig. 3 Scatterplot of university enrollment per capita (left) and the scatterplot of physicians (right), where the y-axis for both represents IPR

4 Regression Models

Having completed exploratory descriptive analysis, we know the data issues that may distort a linear regression model. For model selection, we chose ridge regression, regression trees, and random forest models rather than using OLS, which has strict assumption requirements.

4.1 Ridge Regression

Ridge regression is similar to least squares, but the coefficients are estimated by minimizing a slightly different quantity. More specifically, the ridge regression coefficient estimates $\beta^\wedge R$ are the values that minimize the Eq. (1)

$$\min\left(\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{n}\beta_j x_{ij}\right)^2+a\cdot\sum_{j=1}^{n}\beta_j^2\right),\tag{1}$$

where $a \geq 0$ and is a "panelizing" parameter.

4.2 Regression Trees

The regression tree is a nonlinear model. The algorithm of a regression tree is as follows:

- Use recursive binary splitting to grow a large tree on the training dataset, stopping when all leaves are pure. The minimum number of samples required to split is two, and the minimal number of samples required to be at a leaf node. Samples have equal weight to a leaf node. The maximum number of features is the number of total features in this sample.
- Apply cost complexity pruning (the mean squared error) to the large tree to gain a sequence of best subtrees, as a function of alpha.
- Use k-fold cross-validation to choose alpha. That is, divide the training observations into 10-fold. (a) Repeat Steps 1 and 2 on all but the tenth fold of the training data. (b) Evaluate the mean squared prediction error on the data in the tenth fold, as a function of alpha.
- Average the results for each value of alpha, and pick the parameter to minimize the average error, which is the mean squared error.

$$\min\left(\sum_{m=1}^{|T|}\sum_{i:x\in R_m}\left(y_i-\hat{y}_{R_m}\right)^2+\alpha|T|\right).\tag{2}$$

Here $|T|$ indicates the number of terminal nodes of the tree T, and $\hat{y}_{R_m}$ is the mean of the training observations in R_m, the rectangle corresponding to the mth terminal node.

4.3 Random Forest

The random forest model is an expansion of the tree model. Instead of producing the best tree, we build 100 trees. Each tree is grown based on a subsample generated by bootstrapping technique. The winner is the tree that generates the least mean squared error of all trees.

5 Classification Models

Our next attempt is to use neural network techniques to improve model performance. We first convert IPR into a Boolean variable. Based on IPR's mean and median, any value greater than 2.5 will be coded as 1, meaning a high IPR score, and any value less than 2.5 will be coded as 0, meaning a low IPR score.

After data preparation for artificial neural network (ANN), the first step is to use forward propagation. All input data points are propagated to a single neuron where each input is multiplied with its respective weights and then summed together. Each neuron has an error term or bias. The sum of the bias and the linear combination of inputs and weights is the input to the single neuron on the first hidden layer.

The second step is to apply a nonlinear function, ReLU, sigmoid, or tanh, as an activation function to this linear combination. A proportional neuron on the first hidden layer is randomly selected and contributes to the second hidden layers. After getting the output as a result from forward propagation, the loss can be calculated by the selected loss function. The weights and biases are updated through cross-entropy in order to minimize the loss function.

$$w = w - \frac{\mathrm{d}\varepsilon}{\mathrm{d}w}, \tag{3}$$

where ε is error term and w is the weight. We implement Batch SGD algorithm as the optimizer to update the weight in Eq. (3)

The third step is to use backpropagation to update the weights of the network using the derivative of the cost with respect to a particular weight and shift the value of the weights in that direction. The forward propagation and backward propagation process repeats until the cost function is minimized.

6 Result and Analysis

For the regression modeling, a model performance comparison table is shown in Table 1.

Compared to the ridge and tree models, the random forest regression produces the best outcome, in terms of the lowest RMSE and highest R squares for the testing set. However, the random forest model suffers from overfitting, which can be seen through

Table 1 Regression model comparison result

Regression	Ridge regress (tuned)		Regression trees (tuned)		Random forest	
	Train	Test	Train	Test	Train	Test
RMSE	0.7	0.69	0.626	0.643	0.945	0.534
R squares	0.455	0.423	0.945	0.468	0.945	0.634

Table 2 Neural network model result comparison

Classification	Neural network					
	Model 1		Model 2		Model 3	
	Layer 1 Neurons = 64		Layer 1 Neurons = 128		Lauer 1 Neurons = 258	
	Layer 2 Neurons = 32		Layer 2 Neurons = 84		Layer 2 Neurons = 128	
			Layer 3 Neurons = 32		Layer 3 Neurons = 84	
					Layer 4 Neurons = 32	
Accuracy	0.86		0.85		0.88	
Precision	0.855		0.85		0.88	
Recall	0.855		0.85		0.88	
F1-score	0.86		0.85		0.88	

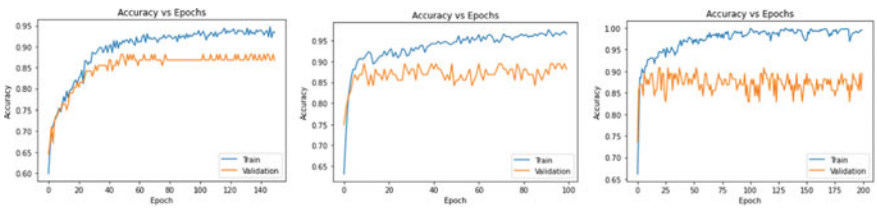


Fig. 4 Model performance in training and validating datasets. The models are ordered 1–3 from left to right

a high R squares score based on the training process. The discrepancy between R squares from training and testing sets is larger than 30% (0.945 − 0.634).

Feature importance extracted based on the random forest model suggests that the top five important attributes in predicting IPR scores are electric power production, GNI per capita, university enrollment per capita, physicians per capita, and percentage annual increase in population.

The results of three neural networks are shown in Table 2.

Based on Table 2, Model 3 seems to perform best, but after examining the model performances in training and validation datasets, Model 1 suffers the least overfitting. Figure 4 displays the details of the model's performance. The first plot, Model 1, shows the smallest gap between the blue (training model) and the orange line (validation model). This indicates model 1 has the least overfitting issue.

7 Conclusion

This article sought to identify the human capital factors with the greatest impact on IPR protections. The resulting models identified university enrollment per capita and physicians per capita to be among the five most important variables in predicting

IPR. These findings support the existing literature and hint at the causal pathways of human capital affecting IPR protections. Human capital has an impact on IPR protections. Future research should build upon these findings to identify the exact causal pathways.

University enrollment per capita was found to be a significant factor in predicting IPR protections. This variable highlights the importance of educational attainment, one of the two primary factors of human capital, in predicting IPR. This identifies that education may be a critical factor driving IPR protections.

The importance of physicians per capita in predicting IPR protections demonstrates the effect of skilled labor. Physicians per capita is an indicator of high human capital and source of R&D in health care. Physicians per capita's predictive importance demonstrates that a high skill workforce is an important element influencing IPR protections.

References

1. Chen Y, Puttitanun T (2005) Intellectual property rights and innovation in developing countries. J Dev Econ 78:474–493
2. Diwakar B, Gilad S (2016) Dynamics of human capital accumulation, IPR policy, and growth
3. Solow RM (1956) A contribution to the theory of economic growth. Q J Econ 70(1)65–94
4. Swan TW (1956) Economic growth and capital accumulation. Econ Rec 32(2):334–361
5. Arrow KJ (1962) The economic implications of learning by doing. Rev Econ Stud 29(3):155–173
6. Romer PM (1986) Increasing returns and long-run growth. J Polit Econ 94(5):1002–1037
7. Romer PM (1990) Endogenous technological change. J Polit Econ 98(5, Part 2)S71–S102
8. Lucas RE (1988) On the mechanics of economic development. J Monet Econ 22(1):3–42
9. O'Donoghue T, Zweimüller J (2004) Patents in a model of endogenous growth. J Econ Growth 9:81–123
10. Romer P (1990) Endogenous growth and technical change. J Polit Econ 98(1990):71–102
11. Gould D, Gruben W (1996) The role of intellectual property rights in economic growth. J Dev Econ 48(2):326–327
12. Falvey R, Foster N, Greenaway D (2006) Intellectual property rights and economic growth. Rev Dev Econ 10:700–719
13. Sattar A, Tahir M (2011) Intellectual property rights and economic growth: evidences from high, middle, and low, income countries. Pak Econ Soc Rev 49(2):163–186
14. Janjua P, Ghulam S (2007) Intellectual property rights and economic growth: the case of middle income developing countries. Pak Dev Rev 46(4):711–722
15. Wozniak GD, Capital H (1987) Information, and the early adoption of new technology. J Hum Resour 22:101–112
16. Riddell WC, Song X (2012) The role of education in technology use and adoption: evidence from the Canadian workplace and employee survey. IZA DP No. 6377
17. Ginarte JC, Park WG (1997) Determinants of patent rights: a cross national study. Res Policy 26:283–301
18. Park W (2008) International patent protection: 1960–2005. Res Policy 37(4):761–766
19. Loukil K (2020) Intellectual property rights, human capital and types of entrepreneurship in emerging and developing countries. Theor Appl Econ XXVII 2020 (622):21–40

Study the Launch Process and Acceleration of a Rear-Wheel Drive Electric Vehicle

Nikolay Pavlov and Diana Dacova

Abstract The paper examines the launch and acceleration process of a rear-wheel drive electric vehicle. An accelerometer and a data acquisition device were used for the purpose of the study. A software program implemented by the authors of the paper makes it possible to study not only the acceleration of the car over time, but also the acceleration as a function of speed and the speed over the time. The paper provides an introduction and analysis of the vehicle launch and acceleration processes. Driving limits determining maximum vehicle acceleration, limited by adhesion between tires on the drive wheels and road are calculated. The used devices are described and compared with the traditionally used in similar studies. The theoretical foundations of signal processing for experimental purposes are given. The program for processing the received data is described. The results were processed and analyzed.

Keywords Electric vehicle · Launch process · Acceleration · Speed · Distance · Numerical integration

1 Introduction

When vehicles move in urban conditions, their motors work for a long time in a transient mode. Frequent launching (starting), accelerating, decelerating, stopping and starting again is required. In this regard, the study of the dynamics of launching and accelerating electric vehicles is a topical issue. For example, in [1] is presented a launching performance comparison between three control strategies under maximum acceleration conditions. The presented experimental data shows a reduction in time when the vehicle is accelerated from 0–60 mph when the launch control system is used. An adaptive launching control strategy of electric vehicle driven by two rear in-wheel motors is proposed in [2]. A MATLAB Simulink computer model of a rear-wheel drive parallel-series plug-in hybrid electric vehicle powertrain as well

N. Pavlov (✉) · D. Dacova
Technical University of Sofia, 8 Kl. Ohridski Blvd, 1756 Sofia, Bulgaria
e-mail: npavlov@tu-sofia.bg

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_29

as a six degree of freedom vehicle model is used to provide a reliable platform for optimizing control strategies of the traction and launch control systems [3]. The launch process when using dry dual-clutch transmission in a conventional vehicle and in a two-speed electric vehicle is studied in [4] and [5], respectively. The vehicle driveline dynamics of launch process is modeled and analyzed in [6]. Acceleration, braking modes and driving cycles of electric vehicles are numerically studied by using MATLAB Simulink in [7, 8]. The accelerations at start-up and deceleration at regenerative braking at different driving modes of a front-wheel drive electric vehicle are studied in [9]. Only one sensor, an accelerometer, is used and after numerical integration of the acceleration time series, driving speed and distance traveled results were obtained.

There are some differences in the starting and acceleration process of front-wheel drive and rear-wheel drive vehicles. In front-wheel drive vehicles, under the action of the inertial force during acceleration, which is directed backwards, the front drive axle is unloaded. This results in a reduction in the ability to transmit traction force between the drive wheels and the road. Launching with intense acceleration can cause slipping and therefore it is not possible to transmit the maximum available torque of the traction electric motor to the road.

In rear-wheel drive cars, the inertial force during acceleration transfers the weight to the rear drive axle and improves the conditions for transmitting maximum torque from the traction electric motor to the drive wheels. In sports cars with high power and torque of the traction electric motor, slippage is possible, although less likely than the front-wheel drive vehicles, in the process of starting and intensive acceleration, and this can also be compensated by using systems for launch control or traction control. The improved traction performance when starting, as well as the good stability during regenerative braking with only an electric motor when driving in a corner of the rear-wheel drive electric vehicles, determines their wide use in modern electric cars built on specially designed new platforms. Such examples are the BMW i3, the new Tesla Model 3, Porsche Taycan, VW ID3 and VW ID4—when they are not in a four-wheel drive variant, only their rear wheels are driven. The return of rear-wheel drive powertrains is observed, which the most automotive companies had abandoned, as the layout constraints and design problems of torque transmission using a traction electric motor are not as big as like the rear-wheel drive conventional ICE cars.

In this paper, the launch and acceleration process of a rear-wheel drive electric vehicle is experimentally studied using a single sensor (accelerometer), and appropriate computational approach to obtain the results for the driving speed and distance traveled is presented.

2 Experimental Equipment and Results

The studies in automotive engineering can be numerical—by using computer programs [2–8] or experimental. Experimental studies can be conducted indoor by test benches [4, 10] and outdoor—on proving grounds or in real road conditions

[1, 9, 11]. When studying the parameters of the vehicle movement—distance traveled, driving speed and acceleration, experimental equipment, called fifth wheel, is commonly used [12–17]. The fifth wheel measures the distance traveled based on the rotation of the wheel. Numerical differentiation of the time series of the traveled distance is used to obtain the time series of the speed and acceleration. The fifth wheel is an expensive device, it is not easy to mount to the car and creates problems for other road users due to the increase in the dimensions of the car.

The easiest way for measuring the vehicle acceleration is with an accelerometer [9], and with the possibilities of modern smartphones which have an accelerometer, even a smartphone can be used for that [11, 18]. We can find the driving speed and the distance traveled using integral calculus. The integral of acceleration over time is change in velocity and the integral of velocity over time is change in position.

The experimental equipment consists of an accelerometer (model 4030-002-120, TE Connectivity) that was mounted on the horizontal surface of the front panel of the electric vehicle. Its measuring range is $\pm$ 2 g, frequency range 0–200 Hz and sensitivity 1000 mV/g. Its sensitive element is capacitive silicon micro-electromechanical (MEMS). The output signal is analog, and an analog-to-digital device (model DQ401, HBM GmbH) is used to convert it for visualization and recording on the hard disk of a mobile computer [9].

The vehicle studied is a pure electric vehicle with permanent magnet motor driving the wheels from rear axle. The main technical parameters of studied electric vehicle are given in Table 1.

Figure 1 shows the forces acting on a vehicle during straight-line acceleration on level road, where G is the vehicle weight, R_{Z1} and R_{Z2}—normal reactions on the front and rear wheels, respectively, F_T—traction force, F_{f1} and F_{f2}—rolling resistance forces, F_a—inertial force, F_W—aerodynamic drag, $\vec{a}$—acceleration vector, h—center of gravity height, L—wheelbase.

The driving limits determining maximum vehicle acceleration, limited by adhesion between tires on the drive wheels and road can be calculated by formula [19]:

Table 1 Electric vehicle technical parameters

Parameter	Symbol	Value	Unit
Vehicle mass	m	1555	kg
Wheelbase	L	2.570	m
Center of gravity height	h	0.470	m
Static load transfer front/rear axle	–	50/50	%
Nominal electric motor power	N	125	kW
Nominal electric motor torque	T	250	Nm
Nominal motor rotational speed	n	11,400	min^{-1}
Number of seats	–	4	–

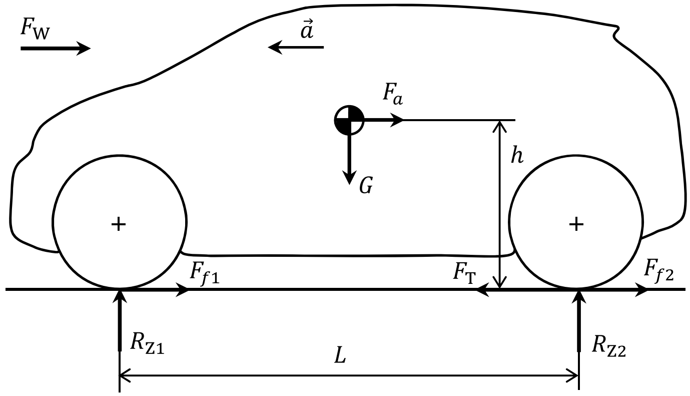


Fig. 1 Forces acting on a vehicle in straight-line acceleration

$$a_\mu = \frac{(D_\mu - f_r)g}{\delta_a} \tag{1}$$

where a_μ is the maximum possible acceleration determined by the adhesion, m/s^2; f_r = 0.012—rolling resistance coefficient; $g = 9.81$—the gravity acceleration, m/s^2; $\delta_a = 1.08$—rotational inertia coefficient [20]. D_μ is the maximum possible dynamic factor for rear-wheel drive vehicle, determined by the adhesion [21]:

$$D_\mu = \frac{(L/2)\mu_x}{L - h(\mu_x + f_r)} \tag{2}$$

where $\mu_x = 0.8$ is the adhesion coefficient between the tires and the road. After calculation with the parameters given in Table 1, the results are: $D_\mu = 0.47$ and $a_\mu = 4.16\ \text{m/s}^2$.

The test was conducted in real road conditions in dry, windless weather on a level road with dry asphalt surface in good condition. Launching was done by pressing the accelerator pedal to the bottom position at the start. In this way, the maximum power characteristics of the electric vehicle are realized right from the beginning of the studied process. The pedal is pressed in fully bottom position until the car reaches its maximum speed. During the test, only the acceleration time series is recorded using the accelerometer mounted on the front panel of the car.

The results obtained during the test are shown in Fig. 2. The maximum realized experimental acceleration is 3.8 m/s^2. The maximum possible acceleration that was obtained theoretically ($a_\mu = 4.16\ \text{m/s}^2$) at an adhesion coefficient of $\mu_x = 0.8$, corresponding to good grip between the tires and the dry asphalt and it is higher than experimental which means that no slippage occurred. The driver also did not notice any signs for slippage. This means that the full power and torque of the traction electric motor has been realized because the accelerator pedal is fully pressed.

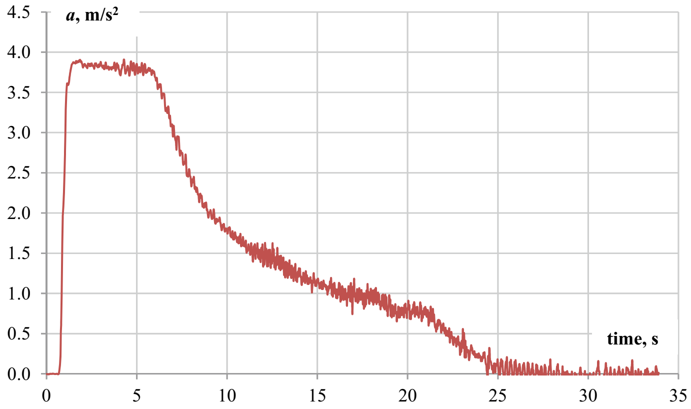


Fig. 2 Acceleration versus time

3 Numerical Calculus

To obtain the other characteristics of the vehicle movement when starting and accelerating, the methods of numerical calculations are used. Numerical integration must be used to obtain the speed. The using of trapezoidal method in MATLAB is possible:

```
speed = cumtrapz(time, acceleration);
speed kmh = speed*3.6;
figure(3)
plot(time, speed_km/h);grid_on, xlabel('time,[s]');
ylabel('speed,_[km/h]')
```

where `acceleration` is the recorded during the test acceleration array, `cumtrapz` is a cumulative trapezoidal numerical integration [22].

Figure 3 shows the result obtained in this way for the speed of the car. From this graph it is easy to determine the time to reach both the maximum speed and speed of 100 km/h. The obtained results for speed can be used to make a graph of acceleration versus vehicle speed (Fig. 4).

After the numerical integration of the speed, the distance is available and plotted relative to the time in Fig. 5.

```
distance=cumtrapz(time, speed);
figure(5)
plot(time, distance);grid on, xlabel('time,[s]');
ylabel('distance, [m]')
```

Figure 6 shows the results for vehicle speed relative to the distance traveled. The time to travel the distance of 0–400 m and the top speed are important indicators for comparing the vehicle dynamics.

The section where the acceleration increases from 0 to its maximum value on the graph of an acceleration versus speed (Fig. 4) as well as on the graph of acceleration

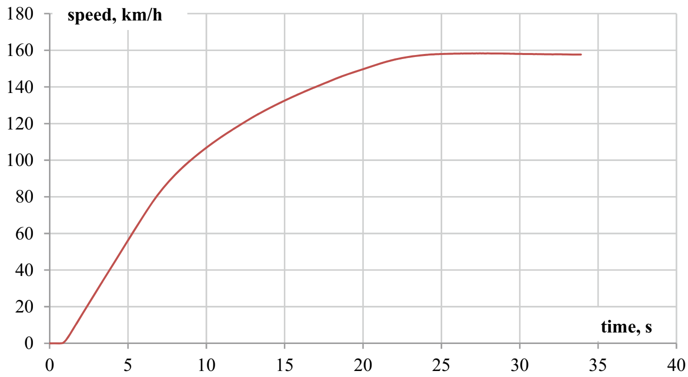


Fig. 3 Speed versus time

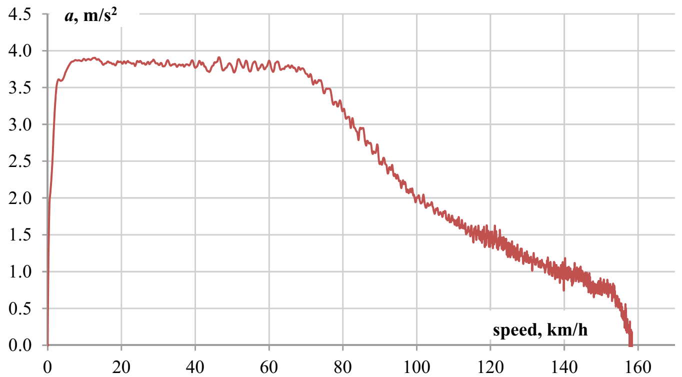


Fig. 4 Acceleration versus speed

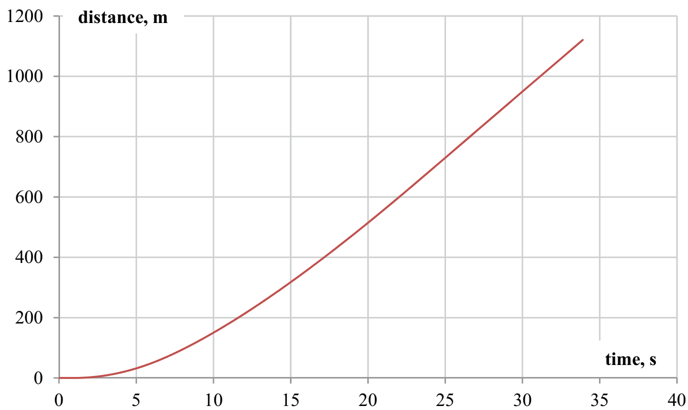


Fig. 5 Distance versus time

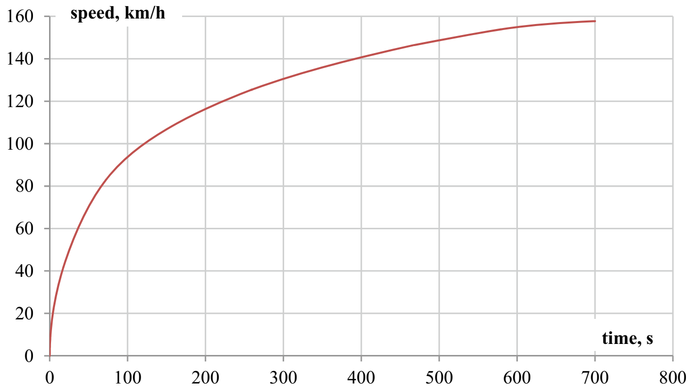


Fig. 6 Speed versus distance

versus time (Fig. 2) give us information about the response of the electric motor to the electric vehicle launch process at fully pressed accelerator pedal. In electric vehicles with lower motor power, as well as in electric vehicles with a higher weight or with a higher rotational inertia coefficient, the time to reach the maximum acceleration is longer.

4 Conclusion

The paper shows the results of test carried out in real road conditions. The conventional methods of studying the launch process and acceleration used fifth wheel or noncontact radar sensors. These devices have a high price and require external mounting on the car body. In this study in-vehicle mounted accelerometer is used. The remaining parameters of the launch process and acceleration to top speed were obtained in software by numerical integration using the trapezoidal method in MATLAB. The results can be used to evaluate and compare the dynamic properties of electric vehicles. Similar software can be developed for a smartphone application. The application can find popularity among drivers to easily test their electric vehicle performances and even diagnosis of the technical condition of a vehicle.

Acknowledgements This research is supported by the Bulgarian Ministry of Education and Science under the National Program “Young Scientists and Postdoctoral Students—2”.

References

1. Compere M, Currier P, Bonderczuk D, Nelson M, Khalifi H (2016) Improving 0–60 mph launching performance of a series hybrid vehicle. Int J Veh Perform 2(3):228–252. https://doi.org/10.1504/IJVP.2016.078558
2. Chen H, Wan Y, Jin D, Zheng S, Lian X (2018) Adaptive launching control strategy of In-wheel motor driven vehicle. In: 37th Chinese control conference (CCC). IEEE, pp 2694–2699. https://doi.org/10.23919/ChiCC.2018.8482584
3. Szechy AM (2016) Traction and launch control for a rear-wheel-drive parallel-series plug-in hybrid electric vehicle. Dissertations and Theses. 315. https://commons.erau.edu/edt/315
4. Zhao Z, Li X, He L, Wu C, Hedrick JK (2017) Estimation of torques transmitted by twin-clutch of dry dual-clutch transmission during vehicle's launching process. IEEE Trans Veh Technol 66(6):4727–4741. https://doi.org/10.1109/TVT.2016.2614833
5. Wu M (2019) Sliding mode control for optimal torque transmission of dry dual clutch assembly of a two-speed electric vehicle during launch. J Phys: Conf Ser 1314:012125. https://doi.org/10.1088/1742-6596/1314/1/012125
6. Sun S, Wu G (2018) Driveline dynamics modelling and analysis of automotive launch process. Int J Veh Perform 4(4):382–402. https://doi.org/10.1504/IJVP.2018.095768
7. Vacheva G, Hinov N, Dimitrov V (2019) Research of acceleration and braking modes of electric vehicles in MATLAB/Simulink. In: 42nd International spring seminar on electronics technology (ISSE). IEEE Press, New York, pp 1–5. https://doi.org/10.1109/ISSE.2019.8810283
8. Hinov N, Punov P, Gilev B, Vacheva G (2021) Model-based estimation of transmission gear ratio for driving energy consumption of an EV. Electronics 10:1530. https://doi.org/10.3390/electronics10131530
9. Dimitrov V, Pavlov N (2021) Study of the starting acceleration and regenerative braking deceleration of an electric vehicle at different driving modes. In: 13th Electrical engineering faculty conference (BulEF). IEEE Press, New York, pp 1–4 https://doi.org/10.1109/BulEF53491.2021.9690780
10. Dimitrov E, Gigov B, Pantchev S, Michaylov Ph, Peychev M (2018) A study of hydrogen fuel impact on compression ignition engine performance. MATEC Web Conf 234:03001. https://doi.org/10.1051/matecconf/2018234001
11. Jiménez D et al (2018) Modelling the effect of driving events on electrical vehicles energy consumption using inertial sensors in smartphones. Energies 11:412. https://doi.org/10.3390/en11020412
12. Peiseler 5th Wheel. https://www.peiseler-gmbh.com/p5rad.html. Last accessed 2022/09/21
13. Pavlov N, Gigov B, Stefanova-Pavlova M (1442) Normative documents for electric vehicles and possibilities for their application in the education of e-powertrain engineers. In: Yilmaz M, Clarke P, Messnarz R, Reiner M (eds) Systems, software and services process improvement. EuroSPI 2021. Communications in computer and information science, vol 1442. Springer, Cham, pp 651–662. https://doi.org/10.1007/978-3-030-85521-5_44
14. Sapundzhiev M, Evtimov I, Ivanov R (2017) Determination of the needed power of an electric motor on the basis of acceleration time of the electric car. IOP Conf Ser: Mater Sci Eng 252:012063. https://doi.org/10.1088/1757-899X/252/1/012063
15. Ivanov R, Sapundzhiev M, Kadikyanov G, Staneva G (2018) Energy characteristics of Citroen Berlingo converted to electric vehicle. Transp Prob 13(3):151–161. https://doi.org/10.20858/tp.2018.13.3.14
16. Ivanov Y, Ivanov R, Kadikyanov G, Staneva G, Danilov I (2019) A study of the fuel consumption of hybrid car Toyota Yaris. Transp Prob 14(1):155–167. https://doi.org/10.21307/tp.2019.14.1.14
17. Kunchev L, Sokolov E, Dimitrov E (2022) Experimental study of transport flows in big cities. In: Proceedings of 21st international scientific conference engineering for rural development 21. Latvia (2022), pp 590–597 https://doi.org/10.22616/ERDev.2022.21.TF192

18. Dacova D, Pavlov N (2021) The study of the possibility to use smartphone in vehicle acceleration measurement. Int Sci J Trans Motauto World 6(3):74–75
19. Dimitrov S, Kunchev L (2016) Theory of the automobile. Publishing House of Technical University of Sofia, Sofia (in Bulgarian)
20. Husain I (2021) Electric and hybrid vehicles. Design fundamentals, 3rd edn. CRC Press
21. Litvinov AS, Farobin YE (1989) Automobile: theory of operational properties. Mashinostroenie, Moscow (in Russian)
22. Chapra SC, Canale RP (2021) Numerical methods for engineers, 8th edn. McGraw-Hill, New York

Measuring Efficacy of the Rural Broadband Initiatives: Evidence from the Housing Market

Hanna Charankevich, Joshua Goldstein, Aritra Halder, and John Pender

Abstract We use proprietary real estate sales data and a variety of empirical methods to account for selection to study the economic impacts of the Broadband Initiatives Program. Broadband Initiatives Program is the largest grant and loan high-speed infrastructure program implemented by USDA and targeted to rural areas. The empirical results suggest that new broadband infrastructure did not have measurable impacts on residential house sale prices.

Keywords Broadband · Rural · Real estate

1 Introduction

Along with water, electricity and transportation, broadband Internet is now an essential part of everyday infrastructure. The importance of a fast and reliable Internet connection became even more apparent during the COVID-19 pandemic, when many

---

The authors thank Stephanie Shipp, Aaron Schroeder, Neil Kattampallil, Anil Rupasingha for their research support and valuable comments and discussions. The authors are also grateful to the USDA Rural Utility Service for their impeccable data work. This project was funded by the USDA Economic Research Service under contract #58-6000-8-00-39.

H. Charankevich (✉) · J. Goldstein
University of Virginia, Arlington, VA, USA
e-mail: hc2cc@virginia.edu

J. Goldstein
e-mail: josh.goldstein@virginia.edu

A. Halder
Drexel University, Philadelphia, PA, USA
e-mail: aritra.halder@drexel.edu

J. Pender
United States Department of Agriculture, Washington, DC, USA
e-mail: john.pender@usda.gov

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_30

people switched to remote work and online distance learning and required broadband to stay connected with their friends and family. As a result, the digital divide between rural and urban areas is a continuing and growing concern in the United States. For example, during 2016 to 2020, on average only 53.6% of rural households had a broadband subscription while in urban areas this number reaches 70.0%.[1]

Broadband Initiatives Program (BIP) was started in 2009 by the Rural Utility Service of the US Department of Agriculture. Its objective was expansion of high-speed Internet in rural areas of the country. In our analysis we are using 79 geocoded BIP services areas merged with real estate transactions data from CoreLogic from 2005 to 2018. This provides us with a decade of property sales both in the pre- and post-BIP period.

Property values are affected by the Internet infrastructure through several channels. Since most communications are Internet-based, education almost always has an online component and remote work becomes more widespread and absence of broadband connection limits demand. Moreover, access to Internet makes it easier for real estate agents to market a house by placing an ad online or conducting virtual open houses. Presence of reliable and secure Internet also eases financial transactions: more offers boost competition and drive housing prices up [1].

The effects of broadband on property prices are becoming a focus of interest of the quickly growing body of literature [1] use census block level data from National Broadband Map to estimate the effects of access to a 25 Mbps connection on average property prices across the country. Using an instrumental variable approach, the authors find that high-speed Internet increases average single-family house price by 3%. However, in the absence of strong instruments these results hardly can be interpreted as causal. While Internet may be a priority for the home buyers in urban areas where most of the population is represented by white collar employees, in the rural areas the effects are not that straightforward. In [2] the author finds positive effects of fiber broadband deployment on property values in rural areas of Germany. In [3] using the same data as in [1], the authors find no correlation of broadband with rural property prices in Oklahoma counties. Though [4] employing a more rigorous triple difference design estimates positive effects of broadband on median house values in rural counties of the United States. They use data on housing outcomes from American Community Survey and broadband availability from Federal Communications Committee Form 447 at the county level.

Unlike previous literature, in this paper we focus on geographical expansion of broadband infrastructure in the rural areas of the United States rather than Internet speed and we pinpoint prices down to the individual property sales. To empirically estimate the effects of Broadband Initiatives Program we leverage two sources of variation: (i) geographical variation in property location inside or in the vicinity of the program service area and (ii) time variation in the program implementation dates. The broadband effects then are recovered using a difference-in-difference

[1] The percentages are obtained using American Community Survey five-year estimates of 2016–2020.

framework by comparing property prices inside the BIP services areas with property prices within 10 mi radius of its border in pre- and post-program time periods.

A major concern for our empirical strategy is that the properties in rural areas selected for Broadband Initiatives Program and properties in areas that chose not to apply to the program or were not selected or ineligible to participate are systemically different and these unobserved differences are correlated with the property prices. We are able to address these issues by using a sample of properties matched on observable characteristics using a Mahanolobis distance [5] to construct a more comparable control group. Then, our second set of estimates is obtained as difference-in-differences on the matched property sales. Moreover, we show that similar results are obtained by utilizing a matching estimation [6] technique that does not rely on panel data variation for identification.

Our results demonstrate that home prices are, in fact, not affected by the new broadband infrastructure in the area. Across all estimation techniques, the estimated effects of Broadband Initiatives Program remain negligible and insignificant. Various robustness checks show that these results cannot be attributed to anticipatory or spillover of the program effects on the properties in the immediate neighborhood of the program service areas. We also did not find differences in the impacts on prices across broadband technologies or size of the program award. Even though BIP was directed to rural areas the properties in these areas did not benefit more from broadband infrastructure than properties urban areas.

Section 2 described the USDA Broadband Initiatives Program. Section 3 details out data and sources and Sect. 4 outlines the empirical specifications. Section 5 reports the results. Section 6 concludes.

2 Broadband Initiatives Program: Background

The Broadband Initiatives Program (BIP) was established in 2009 by the USDA's Rural Utility Service authorized under the Recovery Act. The aim of the program is to improve high-speed broadband access and quality in rural areas. BIP is the largest among USDA broadband development programs in terms of financing: a $2.5 billion appropriation. The BIP provided financing in three forms: grants, loans and grant-loan combination in two rounds of funding in FY 2009 and 2010 to more than 300 projects.

In order to be eligible to participate in the program, at least 75% of proposed project area should be a rural[2] area with insufficient access to high-speed broadband to facilitate economic development.

Definition of the insufficient access to Internet varies by round of applications. In the first round of applications for BIP in 2009 ,only underserved and unserved

[2] Under USDA methodology rural areas are areas not located within a city, town or incorporated area having a population of more than 20,000 and not in an urbanized area that is contiguous and adjacent to a city or town with more than 50,000 population.

areas were considered. Unserved areas were defined as areas in which at least 90% of households lacked access to fixed terrestrial broadband service at a minimum advertised speed of 768 kilobits per second (kbps) downstream and 200 kbps upstream (768/200 kbps). Underserved areas are areas in which at least one of three conditions was met: (1) no more than 50% of households in the proposed service area have access at the 768/200 kbps minimum speed; (2) no broadband service provider advertises service with at least 3 megabits per second (mbps) downstream; or (3) at most 40% of households in the proposed project area have a broadband service subscription.

In the second round the selection criteria were simplified: at least 50% of proposed area must have lacked broadband access at a minimum advertised speed of 5 megabits per second (mbps) (downstream + upstream). In the second round, BIP offered only a standard 75% grant/25% loan combination. All applicants had to propose to provide broadband service to all households and businesses in the proposed area.

The award phase of the second round was completed in September 2010. Under the BIP, RUS awarded $2.2 billion in grants and $1.2 billion in loans to 299 terrestrial broadband infrastructure projects. A total of 63 of the terrestrial infrastructure projects were approved in round 1 and 236 in round 2. Thirty-nine of the projects were not completed and the funds were rescinded. Majority of the funded projects are last-mile infrastructure projects. In addition to funding terrestrial broadband infrastructure, in its second round, BIP provided grants to support satellite infrastructure, broadband in rural libraries and technical assistance. Satellite projects were awarded four grants with a total value of $100 million. And 19 technical assistance grants received about $3 million total.

3 Data

3.1 Housing Data

The source of data on property sale transactions and property characteristics is the national real estate data provider, CoreLogic. The housing data sourced from CoreLogic consists of two parts: sale transactions and tax assessments. Sale transactions contain information on sale price and date, type of transaction and sellers, owners and involved agencies (for example, lending agency and title insurance company). And the tax assessments collect data on property characteristics such as property location, size and age as well as assessed property values for tax purposes. The dataset covers the universe of properties in all 50 states and is collected from publicly available sources such as county appraiser offices and multiple listings service.

The CoreLogic dataset went through multiple steps of preparation. Administrative data records like real estate tax records have, by their nature, multiple entries over time of information related to the same entity. Firstly, we created a unique individual record for each property from multiple administrative record entries. Assuming that the most recent entry is the valid one, we fill in the missing information on property characteristics from the previous records.

Secondly, since for the purposes of our analysis we are interested only in residential properties; the large dataset was initially filtered to derive a sample representing arms-length sales of single-family residences.[3] To construct our sample, we excluded transactions that are labeled "non-arms-length" (transaction type code "9") and kept only transactions referred to as re-sales or sales on newly constructed houses (transaction type codes "1" and "3").

We also want our sample to include only single-family type of properties. This is equivalent to setting the "property indicator" equal to "10". Single-family housing is comprised of single-family residences, townhouses, apartments, condos, co-ops, flats, multiplexes, row, mobile, manufactured homes, etc.

CoreLogic collects extensive data on property characteristics and includes more than 200 variables. However, many of those are not complete or available only for a certain type of properties. For our analysis, we consider the following property characteristics: location, sale information, size and age, number of bedrooms and bathrooms. Property location is described either by a full property address or by parcel-level centroid geographical coordinates. Sale information includes latest sale amount in dollars and the date when the sale contract was signed. We drop any sales prior to 2005. Size of residence is represented by either a building size in sq. ft. or a size of the living area in sq. ft. Land square footage or property acreage is used to describe size of property. The construction year of the original building or the year when the building was first assessed with current components is used to calculate the age of the property. We used the Fannie Mae and Freddie Mac Uniform Appraisal Dataset Specification to calculate number of baths.

Additionally, each property in the selected sample was assigned a census block identifier that allows us associate demographic data collected by the census surveys with each property. We create these identifiers by finding the geographic intersection of the census block areas with the geographical coordinates of the CoreLogic properties.

Finally, to ensure the validity of each observation we eliminate clear mis-entries in the dataset. For example, a 1200 ft^2 house with 75 bedrooms. The mis-entries are detected as multi-attribute outliers using Cook's Distance.[4] The final sample of relevant residential property sales consists of more than 20 mln transactions over the period of 2005–2018.

---

[3] Arms-length sales transaction (primary sale code "A") is what we might call a "typical" transaction between two parties, not a special transaction between parties, such as a sale to a relative for a reduced amount.

[4] Cook's Distance is an estimate of the influence of a data point. It takes into account both the leverage and residual of each observation. Cook's Distance is a summary of how much a regression model changes when the *i*-th observation is removed.

3.2 Broadband Initiatives Program Data

We obtained data on Broadband Initiatives Program from the Economic Research Service, US Department of Agriculture. The BIP data consists of two files. Both files can be linked together by a unique program ID. The first file is a collection of shape files describing geography of each project award. Shape files are geographical boundaries of funded areas. Shape files define the funded areas either as a map polygon for a single-area project or as a multi-polygon for projects covering multiple areas. We were able to recover full shape files for 228 BIP projects. The number of sub-projects is 1168.

The second file contains the project area and award characteristics. These characteristics describe whether a project is last mile or middle mile, broadband technology used, as well as how many households and businesses are in the project area and how large is the program award amount.

To match the BIP program data to the data on property transactions, we expanded projects' geographical borders in the shape files by 20 miles like pictured in Fig. 1 (a solid outer line). Then we use the parcel level longitude and latitude to find out which of the properties are located within the program's boundaries and inside the expanded boundary. We consider all the properties that fall within 20-mile distance from the programs boundaries as properties not impacted by the BIP. These properties comprise a control group. In Fig. 1 these properties are pictured in orange. And all properties that are located inside the BIP boundaries are considered as treated properties, as prices on these properties can be affected by the broadband expansion into the area. These are the green dots in Fig. 1.

The final analytical sample includes 108 projects with non-zero amount of property sales in the selected area and 2.7 million property sales. A total of 2.6 million of these transactions fall outside the project boundary but within 20-mile distance of it and more than 124 thousand are inside the BIP boundaries. We were able to match project information for 79 BIP service areas. The descriptive statistics of these projects are in Table 1. Average number of households in the BIP service area is about 10 thousand and average number of businesses is 1.4 thousand establishments. Some project areas do not include any of the business entities. The smallest amount per household is 57 dollars while on average BIP award was 4 thousand dollar per household in a service area. Fiber technology (FTTH) is the most frequently provided technology by the BIP.

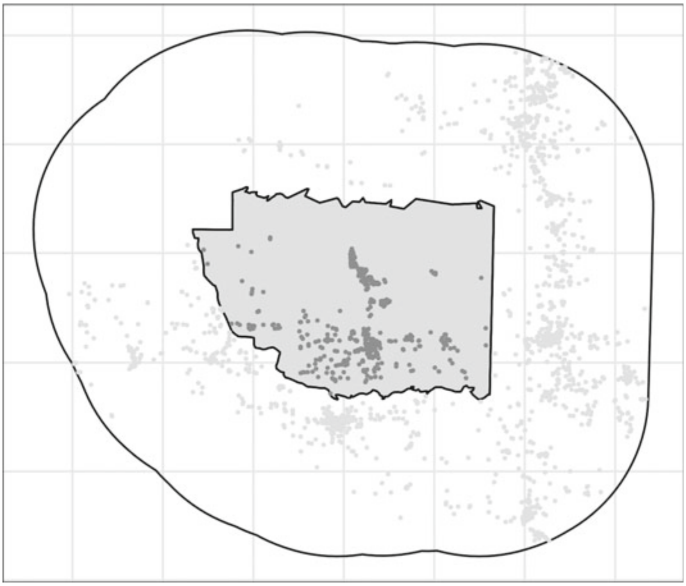


Fig. 1 Example of BIP service area with an extended 20 mi border (**a**) Property sales (**b**) Inside (dark grey) service area and outside (in light grey). The figure pictures one of the BIP project service area shapes located in Oklahoma

Table 1 Descriptive statistics of broadband initiatives projects selected for analysis

	Min	Mean	SD	Max
No. households	352	10,507.15	19,332.91	105,904
No. businesses	0	1416.28	2903.14	18,621
Award per HH (USD)	57.3	4009.13	3406.08	15,664.98
FTTH	0	0.608	0.491	1
Wireless	0	0.329	0.473	1
DSL	0	0.152	0.361	1
Observations	79			

4 Empirical Methods

4.1 Difference-in-Difference Estimation

We are interested in evaluating the average effects of the Broadband Initiatives Program on property prices. In other words, we want to estimate how the property prices would have had changed in the absence of broadband in comparison with prices we observe after program implementation. However, the counterfactual outcome is not observed, and we rely on the observational data of properties in the

20 miles radius of the BIP area. Given a non-randomized nature of the program and panel data on property sales, we employ the two-way fixed effects (difference-in-difference) design, which compares the changes in property prices for residences in the BIP areas before and after the program implementation to changes in prices before and after for properties in the neighborhood of the BIP program area.

For estimation purposes we divided the analytical sample into six two-year periods: 2005–06, 2007–08, 2009–10, 2011–12, 2013–14 and 2015+. This includes three pre-program periods and three post-program periods and also helps us to test for the existence of pre-program implementation trends. The 2005–06 is the base period. The main specification takes the following form:

$$\log\left(P_{ijt}\right) = \alpha_0 + \alpha_1 \mathrm{BIP}_i + \sum_{j=1}^{J=5} \beta_j \mathrm{Year}_j + \sum_{j=1}^{J=5} \gamma_j \mathrm{BIP}_i \times \mathrm{Year}_j + \mathrm{Tract}_t + \varepsilon_{ijt}, \tag{1}$$

where i indexes property, j—time period and t—census tract. $\log(P_{ijt})$ is natural logarithm of the sale price of property i in time period j in census tract t. BIP_i is the BIP indicator variable that takes the value of 1 if property i is located inside the program area. Year_j are the time periods dummies where $j = 1$ corresponds to the 2007–08 and $j = 5$ indicates 2015+. Tract-level fixed effects Tract_t accounts for time invariant differences between tracts in analysis area such as rurality and local policies, ε_{ijt} is the error term, coefficients of interest γ_j capture the program effects in time periods $j \in [1, 5]$ relative to the base period of 2005–06.

To account for possible confounding effects in our regression analysis we control for property and program characteristics. Property characteristics include sale transaction type, number of bedrooms and bathrooms, size of land and living area and ratio of living area to the total building area and age of property measured as the difference between the sale year and year the property was built or lastly modified (effective year built). BIP characteristics control for different broadband technologies and size of award per household. We differentiate three technology types: FTTH, wireless and DSL (asymmetrical and very-high speed DSL). FTTH technology may be FTTH GPON, FTTH RFOG or FTTH PTP. Wireless means both fixed wireless and mobile wireless technology. We dropped eight projects that provide power line and hybrid fiber-coaxial cable technology.

4.2 Mahalanobis Matching

To address the concern of covariate imbalance, we implement a matching algorithm to find properties outside the BIP that look similar to properties within the BIP. We match properties inside the BIP area one-to-one on properties within a 0–10, 5–15 and 10–20-mile radius of each BIP area. We require exact matches by project and in the number of bedrooms and bathrooms. Then, the Mahalanobis metric is used

Table 2 Balance on matching covariates in matched property sales

	Outside	Inside	*p*-value
Age	33.249	33.079	0.215
Bedrooms	1.946	1.946	1
Bathrooms	1.395	1.395	1
Living/building sq. ft	1.249	1.253	0.014
Land sq. ft	66,105.47	110,614.6	0
Living sq. ft	1834.674	1845.159	0.005
Observations	94,236	94,236	

Note The table presents the results of two-sample unpaired *t*-test. The first column is mean value among property sales outside program boundary and second column—among properties inside the project service area. The third column is the *p*-value the *t*-test. Variance is assumed to be equal

for 1:1 nearest-neighbor matching to identify the property that is most similar in the remaining property characteristics including lot size, living area, ratio of living area to building area and effective year built. The Mahalanobis distance between property i inside the BIP and property j outside is defined as

$$\delta(\boldsymbol{X}_i, \boldsymbol{X}_j) = \sqrt{(\boldsymbol{X}_i - \boldsymbol{X}_{\mathrm{j}})' S^{-1} (\boldsymbol{X}_i - \boldsymbol{X}_j)}$$

where $\mathbf{X}$ is a vector of matching covariates and S is the pooled covariance matrix.

We assess the ability of our matching algorithm to produce balanced samples by comparing mean differences in standardized values across these covariates in properties inside and outside of the BIP area. Table 2 gives the balance in matching covariates after matching for properties within 10 mi of the BIP border and inside the service area. We cannot reject the equivalence in average values at 5% significance level in all matched property characteristics but size of the living area and land.

The program effects then are estimated applying a bias-corrected estimator as in [6] to each year of observations in the matched sample and comparing the coefficients on BIP dummy between years preceding program and post-program implementation.

5 Results

In this section we review the results of empirical analysis of the BIP effects on property. Table 3 presets the results from estimating Eq. (1) using sample of properties within 10 mi of the program border as a control group. The results using sample of properties in 15 and 20 mi radius remain the same. We focus our discussion on the specification that includes property and program controls and census tract level fixed effects as the estimates were similar in magnitude and significance for specification with and without covariates and/or fixed effects. Column 1 reports the estimated coefficients α_1, β_j and γ_j using the whole sample. Column 2 reports same coefficients but estimated using a sample of matched properties [7]. The estimates suggests that on overall BIP has a positive effect on property prices, however these effects dissipate as we add time trends. Coefficients γ_j are negative and close to zero. This pattern becomes even more pronounced in the matched sample: the estimated effects of BIP vary between − 0.1% and − 0.4% and are significant only for time period of 2011–12. It is implausible that benefits of the new broadband infrastructure have already propagated into the property prices in the two years following the program award, but it may reflect the inconveniences associated with construction works in the program area.

Estimates of BIP effects obtained using matching estimator reported in Table 4 paint a similar picture. We find no significant improvement in housing prices after program in comparison with estimates in 2005–06 years period. The estimated coefficients in post-program years still indicate that property prices in the program area are lower than in the 10 mi neighborhood.

Figure 2 compares the estimated effects of BIP of property prices in the 10 mi radius from the program area border using three estimation techniques described in previous section. The figure reveals that the estimates program effects are not statistically different from zero. Econometric analysis of the property prices suggests that the BIP grant program had no substantial effect in increasing residential house values.

Table 3 Estimated ITT effects of BIP: main results

	Estimate	Std. error	Estimate	Std. error
	(1)		(2)	
	DID fixed effects		Fixed effects on matched	
BIP	0.019	(0.008)	0.004	(0.001)
Year 2007–8	0.013	(0.007)	0.002	(0.001)
Year 2009–10	− 0.14	(0.007)	− 0.011	(0.001)
Year 2011–12	− 0.201	(0.007)	− 0.016	(0.001)
Year 2013–14	− 0.095	(0.006)	− 0.009	(0.001)
Year 2015 +	0.055	(0.005)	0.005	(0.001)
BIP × Year 2007–8	0.034	(0.011)	− 0.007	(0.008)
BIP × Year 2009–10	− 0.008	(0.012)	− 0.001	(0.002)
BIP × Year 2011–12	− 0.055	(0.011)	− 0.001	(0.002)
BI P × Year 2013–14	− 0.041	(0.010)	− 0.004	(0.001)
BIP × Year 2015+	− 0.026	(0.009)	− 0.001	(0.001)
Covariates	Yes		Yes	
Tract FEs	Yes		Yes	
Observations	173,680		135,189	

Note Standard errors are in parenthesis. Year 2005–6 is the omitted category

Table 4 Estimated ITT effects of BIP: matching results

	Matching		
	Estimate	Std. error	Observations
$BIP_{2006–05}$	− 0.038	(0.007)	22,468
$BIP_{2007–08}$	− 0.022	(0.008)	19,072
$BIP_{2009–10}$	0.006	(0.008)	17,994
$BIP_{2011–12}$	0.003	(0.008)	23,366
$BIP_{2013–14}$	− 0.002	(0.006)	33,418
BIP_{2015+}	− 0.023	(0.004)	72,154

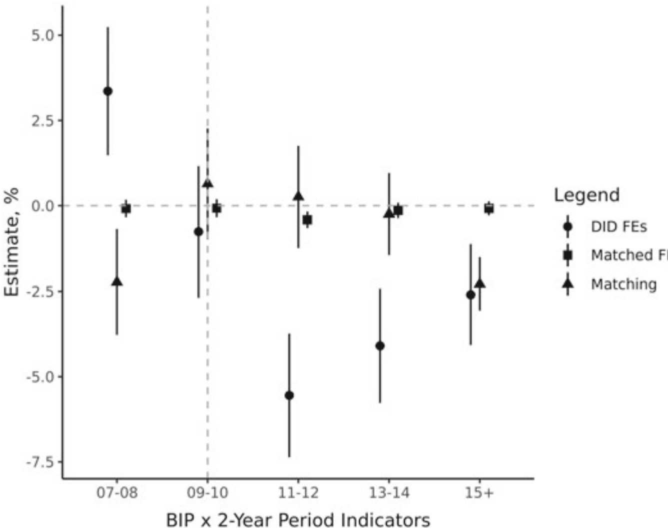


Fig. 2 Estimated effects of BIP on property prices using DID with fixed effects, DID on matched sample and matching. The upper and lower whiskers represent the 95% and 5% confidence intervals, respectively

6 Conclusion

High-speed reliable Internet connection has become increasingly important in the last years. Broadband development in sparsely populated geographic areas such as rural areas has been facilitated by various government-funded programs. In this paper, we are using proprietary housing dataset to quantify the efficacy of Broadband Initiatives Program (BIP) through property prices in rural areas. Using a combination of methods, we estimate the impacts of new broadband infrastructure to be negligible and negative in 2, 4 and 5+ years post-program implementation. These results remained insignificant after we addressed for the selection bias using matching on observables. One of the drawbacks of our analysis is that we do not have information on the status of broadband adoption at the household level in the service areas: we are not able to check whether houses in the program areas are now connected to broadband Internet or whether the Internet speed has increased.

References

1. Molnar G, Savage SJ, Sicker DC (2019) High-speed Internet access and housing values. Appl Econ 51(55):5923–5936
2. Klein GJ (2022) Fiber-broadband-internet and its regional impact—an empirical investigation. Telecommun Policy 46(5):102–331
3. Conley KL, Whitacre BE (2020) Home is where the internet is? High-speed internet's impact on rural housing values. Int Reg Sci Rev 43(5):501–530
4. Deller S, Whitacre BE (2019) Broadband's relationship to rural housing values. Pap Reg Sci 98(5):2135–2156
5. Gu XS, Rosenbaum PR (1993) Comparison of multivariate matching methods: structures, distances, and algorithms. J Comput Graph Stat 2(4):405–420
6. Abadie A, Imbens GW (2011) Bias-corrected matching estimators for average treatment effects. J Bus Econ Stat 29(1):1–11
7. Stuart EA, Huskamp HA, Duckworth K, Simmons J, Song Z, Chernew ME, Barry CL (2014) Using propensity scores in difference-in-differences models to estimate the effects of a policy change. Health Serv Outcomes Res Method 14(4):166–218

Critical Junctures in Contemporary Media and Communication Processes (Bulgarian Case Study 2000–2020)

Lilia Raycheva, Bissera Zankova, Nadezhda Miteva, Neli Velinova, and Lora Metanova

Abstract The case study provides analysis of some of the critical junctures in Bulgarian media ecosystem developments, based on the research of major sources and datasets on media and journalism in the country (2000–2020) in four domains (legal and ethical regulation, journalism, media usage patterns, and media-related competences). While in some of the domains research is well presented; in others it is not comprehensively developed. Comparatively well advanced are the analyses of media-related legislation and regulation domain, although empirical practices are less explored. Reasoning on the media structure developments is more thoroughly approached in view of freedom of expression, freedom of information, and the ethical issues of media accountability. The journalism domain is addressed through market developments, public service media, content production, and work conditions. Media usage patterns are examined with the prevalence of issues regarding pluralism of viewpoints, relevance of news media, and trust in media. The domain of media-related competences is of growing scholarly interest, especially in the area of media literacy initiatives and sustenance of professional standards. The analysis of the selected sources supplements tracking the critical junctures between the various elements of deliberative communication, which provides ground for outlining the perspectives of the media developments in the country.

Keywords Deliberative communication · Media · Critical junctures

L. Raycheva (✉) · N. Miteva · N. Velinova · L. Metanova
The St. Kliment Ohridski Sofia University, Sofia, Bulgaria
e-mail: lraycheva@yahoo.com

B. Zankova
"Media 21" Foundation, Sofia, Bulgaria

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_31

1 Introduction

During the first two decades of the twenty-first century, the Bulgarian media ecosystem has experienced intensive processes of transformation, impacted by the vigorous information and communication technologies, which were supplemented by new economic models of production and consumption of media content.

That is why, in the considered 20-year period, the media research interest in the country has been focused primarily on the challenges of the political, economic, social, and professional aspects of digitalization, reflected in various aspects of media developments, such as: legislative and regulatory; journalistic practices; media usage patterns; and media-related competences, all covered by the MEDIADELCOM project "Finding risks and opportunities for European media landscapes." It started in March 2021 as a three-year research project financed by Horizon 2020—the EU funding program for research and innovation. MEDIADELCOM involves 17 teams representing 14 EU countries from Western, Central, and Eastern Europe, and it is coordinated by the University of Tartu (Estonia) [1].

In particular, legal framework, regulatory practices, and civil ethical initiatives in Bulgaria are comparatively and comprehensively studied at national and international level. Despite the publications related to the topic data about the Bulgarian media regulation have been also collected through other European projects and surveys submitted by the national ministries to the CoE, EC, OSCE, UNESCO, or ITU. In the research international principles and aspects of freedom of expression and freedom of access to information as well as the acceptable limits of these fundamental rights dominate.

The resulting changes in the nature of the journalistic profession, the role of the media, and journalists in the digitalized socio-economic conditions are also comparatively well researched. Regarding the quality of the media content, the following main characteristics have been studied, although sporadically: timeliness of the news programs; public significance of the broadcast information; factual accuracy based on verification by independent sources of information; objectivity—disclosure of all facts in an unbiased way; presentation of plural points of view on the topic; publication of in-depth journalistic works on socially significant topics (investigations, reports, analyses, comments); writing and spelling style; etc. Along with many benefits and positive effects of the new media ecosystem, increasing trends to misinformation, manipulation, and hate speech have also been examined.

The media usage by audiences is studied in light of several factors such as access to media content, media diversity, functionality and quality of the media, public trust in the media, and new media. The most common research is related to public trust in the media and frequency of media consumption, broken down by different age groups, as well as divided into social and ethnic principles. The type of media preferences (TV, radio, print, internet, websites, social networks, and social media) has been also studied, as well as variety of issues, regarding media consumption and quality of news content.

Research on media-related competencies is rather sporadic. Specific interest especially on media literacy issues has been growing lately, mainly due to the efforts of non-governmental organizations and academia.

Following the aim to emphasize on the state of the art of the existing research in the country with regard to risks and opportunities for deliberative communication, a large array of specialized publications has been identified and examined. This includes predominantly findings of transnational organizations that monitor media systems globally; datasets of national statistics and public bodies; legislative, policy, and regulatory documents; institutional official papers and non-government reports; academic national and international research; major sociological surveys; research of non-governmental organizations; publications of professional media associations, etc., using keywords related to the four domains. Large comparative research projects that collect periodically data and produce comparative analysis over certain periods are relatively scarce and inconsistent, as well as thorough commentaries of the media industry.

Some of the entities engaged with the provision of documents and expert positions, relevant for all domains, are: The Union of Bulgarian Journalists; The Council for Electronic Media; The Communications Regulation Commission; The Ministry of Transport, Information Technologies and Communications; The Ministry of Culture; The Bulgarian Association of Communication Agencies; The Branch Association of Bulgarian Telecommunication Operators; universities; The Bulgarian Academy of Sciences; The Bulgarian National Television (BNT); The Bulgarian National Radio (BNR); providers of media services; The Konrad Adenauer Foundation (KAS); The Open Society Institute-Sofia; The Reuters Institute; Reporters without Borders; Freedom House; The National Council for Journalistic Ethics Foundation; The Access to Information Program NGO, etc.

Significant journals dealing with media-related issues are predominantly distributed online: Rhetoric and Communications; Newmedia21; Media and Public Relations; Postmodernism Problems; etc.

The National Statistical Institute provides substantial statistical data about some activities of the press organizations and the audiovisual media providers.

2 Political and Social Changes Outlining the Trends in the Bulgarian Media Developments

The transition from one-party political system and centralized economy to democratic and market forms of government and economy after the socio-economic changes in the country of 1989 lasted for a long time. Only in 2002 in its annual report the European Commission recognized Bulgaria as a country with a functioning market economy [2]. In 2004, it became a member of NATO, a necessary condition for all former socialist countries to join the European Union. In 2007 the country's membership in the European Union became a reality.

Initially, the changes in the Bulgarian mass media system and the directions for its development were interrelated with the political, economic, and social dynamics in the country. The processes of demonopolization, decentralization, and liberalization were formed arbitrarily laying the foundation for building the new media environment. These processes were accompanied by a general shortage of financial, technological, and human resources to be mobilized and concentrated in the service of the current priorities of change, based on the values of civil society and the mechanisms of the market economy.

Researchers underline the fact that the creation of the democratic Bulgarian media system has taken place chaotically and without clear rules and frameworks. The reform in media policy, regulation, and accountability is characterized as being slow, "while the steps taken towards state emancipation, liberalization and privatization were overhasty, unpremeditated and premature. The consequences of that approach were that strategic economic and political allegiances have started exerting serious power over media content through direct editorial control, gate-keeping of information, bias in representation, programme choice, commercialization and the tabloidization of press and electronic media formats towards more entertainment, sensationalism and scandallousness" [3]. The processes of demonopolization, decentralization, and liberalization were inconsistent [4]. The lack of a national concept and strategy for the transitional development of the Bulgarian media environment turned out to be among the extremely important reasons for its incomplete transformation [5]. The systematic approach was missing, regulation was delayed, and the pursuit of rapid profits in this area prevailed over the public interest. The gloomy observation is that "in the absence of clear normative standards media is increasingly seen as extension of either partisan or corporate strategies" [6]. Thus the transformation of the Bulgarian media system has been premised on political and commercial interests and not on public values.

These deficits laid the basis for the shortcomings in media maturing, noted in the 2013 initiative of the Open Society Foundation for studying of digital media in 60 countries. Among the problem areas were the frail media legislation and regulation, the lack of energetic institutional measures against media concentration; the uncontrolled media consolidation; the departure from professional standards; the lack of pluralism of opinions and diversity of content, etc. The positive aspects were outlined mainly around the "activities of the civil society, which in specific cases had clear impacts on both politics and commercial media" [7].

The reasons for these shortcomings are complex. Particularly media property and media concentration have never been dealt with properly through an adequate and transparent regulatory framework. On the other hand, the attitude of journalists toward non-transparent media ownership and the distribution of print publications according to a study "Journalism without Masks" carried out by the Association of European Journalists—Bulgaria (AEJ-Bulgaria) and Alpha Research Sociological Agency has remained unchanged since 2015. It pointed that this is a problem of ultimate importance—for journalists and for the future of the media. Every second respondent noted that regulating media ownership and cross ownership is the first measure that should be applied to improve the media environment in the country (55%) [8].

During the first two decades of the twenty-first century, the transformation processes in the Bulgarian media ecosystem were intensified, due to the impacts of the digital technologies and the new economic models of production, dissemination, and consumption of media content. These technologies improved the means and the ways of communication, which catalyzed both the horizontal exchange of information between people living in one and the same period of time and its vertical transmission to offspring. However, the media environment became much more complicated, and problems in it were augmented. These processes were taking place against the background of the still unfinished transition from a full state monopoly to diversification of the media and their functioning in market conditions. Remnants of this monopoly can be clearly seen today in the mechanisms of financing the public service media, which questions the independence of the Bulgarian National Radio (BNR) and Bulgarian National Television (BNT) from the ruling political class. The big commercial media are also dependent on the governments and fight for their favor in the financial disbursement for media companies, provided by EU funded programs. As reported by the Open Society Institute in 2005 and in 2008 political elites remained determined to keep public service broadcasters under tight control after the democratic changes, and this took place with a greater or lesser intensity across the Central and Eastern European region. When these countries became members of the Council of Europe and later acceded to the EU, it was critical that they should meet existing European standards of public media independence. During the period of negotiation before entry politicians refrained from influencing public service media [9]. Thus a critical merge of politics, business and media threatened freedom of expression and freedom of the media. Deregulation of the radio and television broadcasting sector was protracted, giving way to the rise of two interrelated processes—politicization of media and mediatization of politics [10]. Since the beginning of the new century, these processes have accelerated with the widespread use of digital technologies in everyday communication. It is notable, though, that, according to the World Press Freedom Index, while in 2006—the year before accession to the European Union, Bulgaria ranked 36th, while in 2020 it collapsed to 112th place among 180 countries in the world [11].

In 2021, Bulgaria ranked among the Member States of the European Union with an average level of digitalization [12]. According to EU's Digital Intensity Index 2021 Bulgarian business had the lowest level in the EU in digitalization and investment in digital technologies [13]. This certainly did not apply to the major media and telecommunications companies in Bulgaria. However, the country is still experiencing significant delays and difficulties in building an e-government to consolidate e-data and services for the benefit of businesses and citizens.

Data provided by the National Statistical Institute present the trends for the media developments in the country. The decrease in titles and circulation in print media is notable: In 2020 there were 209 newspapers with annual circulation of 123, 287 mln copies (dailies—33; published 2–3 times a week—11; published less than once a week—71; and weeklies—94). In comparison, prior to the EU accession in 2007 there were 423 newspapers on the market with annual circulation of 310, 023 million copies. Radio stations and TV channels mark decline versus increase in hours broadcasted. While in 2007 there were 222 national television channels with 599, 135 h of programming and 150 radio stations with 591, 836 h of programming, in 2020 they were reduced to 120 TV channels (779, 830 h.) and 77 radio stations (635, 102 h.). On the contrary, Internet penetration for households in the country has increased more than four times for the same period: 17.0% (2007) to 78.9% (2020) [14].

Despite the rapid development of ICT and online services, television continues to be the most preferred source of information and entertainment for most Bulgarian households. In addition to traditional media and online-only news sites, using of other social media platforms, as well as networking and microblogging services such as Facebook, Google Plus, Instagram, Twitter, TikTok, and hashtags, is becoming more and more popular. The use of online social networks every day or almost every day is 56% (in EU it ranges from 46% in Germany and France to 77% in Lithuania) [15]. The creative potentials of the new information and communication environment appear to be a key factor in the development of the Bulgarian media reality. More than 76% of the Bulgarians use Facebook for any purpose and 64% for news; 70/64%—YouTube; 54/17%—Facebook Messenger; 61/16%—Viber; 36/12% Instagram; and 13/8%—Twitter. About 38% share news via social media, messaging, or email [16].

3 Risks and Opportunities of Media Developments

The review of the existing research and the conducted analysis of the media environment in Bulgaria allows highlighting the critical junctures in the media environment in the country in the period 2000–2021.

Although the country is defined as free in terms of political and civil rights (Freedom House) [17], freedom of expression (Reporters without Borders), journalism and media market are at increasing risk of instability and dependence. The freedoms of movement of goods, capital, services, and people of the European single market turned to be challenging to upholding of the basic pillars of Europe's audiovisual model, such as cultural diversity, media pluralism, and protection of minors, consumer protection, and intolerance of incitement to hatred.

3.1 Legal and Ethical Regulation

The selected sources in legal and ethical regulation domain present the results of in-depth research on media law and media regulation of radio and television environment and the main aspects in self-regulation and media ethics. They also cover the legal framework of digitalization of the electronic media and the main regulatory ideas concerning the new online media. The opportunities and challenges generated by new media services for media freedom and independence are also examined. Possible critical junctures may arise as a result of the slow and incomplete media legislation, non-systematic implementation, and the deficiency of media accountability, media self-regulation, and media co-regulation. The lack of strong and demanding civil society, constant political pressure, and submissive journalistic culture which does not vie for independence and high moral standards every day are other factors that have to be taken into consideration. A critical juncture could arise from the upcoming application of the EU digital services and digital markets package which will require close cooperation and harmonization of the actions of Member States in the complex digital environment to enable transparency, user safety, and platform accountability against the trade of illegal goods, services, and content online and manipulative algorithmic systems spreading disinformation [18].

3.2 Journalism

The examination of the selected research sources on the media environment in Bulgaria (2000–2020) draws attention to several critical junctures for Bulgarian journalism. Most of them are related to media pluralism in its various aspects—diversity of content and opinions, transparency of media ownership, political and financial (in)dependence of the media, social exclusion of groups from society, etc. The state of the journalistic profession in the market and work conditions, education and realization of students in journalism, and journalistic values and standards reveal additional risks for the development of journalism in the country.

All forms of media pluralism are threatened, the most critical being the state of market pluralism, political, and corporate independence of the media. Other serious problems are commercialization of journalism, deterioration of the working environment and labor market for journalists, lowering professional standards, declining consumer trust in traditional media, and the rise of online platforms. Opportunities to improve the media environment stem from overcoming the risks themselves. They require the will and coordinated action of political class, legislature, media owners, media and communication regulators, professional journalistic organizations, academia, and civil society.

3.3 Media Usage Patterns

The analysis of the research regarding media usage patterns shows that although considerable amount of reliable data is available, it is not sufficiently regular and systematic. Two main critical junctures can be outlined in this domain: the decline of public trust in media due to their economic and political dependence and media consumption divide by age and social groups due to technological developments.

Audiences increasingly prefer easily digestible information, preferably presented through video. More and more people are relying on social media to choose information, becoming more and more inert in their search for media. The leading device for reading and watching news is the smartphone, which is decisively ahead of the personal computer. A different approach of the young generation to media consumption has been noted. The center of gravity of young people has shifted from the professional sphere to leisure; consumption for younger generation is often a more important identifier than career or status. Thus, the public is displaced by the private; the communities—by the networks. The adult population between the ages of 66 and 75 are heavy users of traditional media—television, press, and radio [19].

3.4 Media-Related Competences

In the analyzed sources several critical junctures with regard to media competencies stand out. They are connected with trainings to increase media and digital literacy; coping with fake news and misinformation; media diet preferences; and technological challenges. After the COVID-19 pandemic, when everyday life moved online and even older people who had not actively used the Internet and social networks had changed their habits, it became clear that they also needed media literacy. Topics such as how to distinguish reliable from unreliable sources of information; how to recognize fake profiles on social networks; how to protect oneself from online fraud; what are the risks associated with one's personal data online; and how to select sources of information are challenging to the broader audiences.

Media and digital literacy are perceived as an effective remedy against the spread of fake news and misinformation, as a tool for creating and training of critical and analytical thinking. The fact that more than half of the Bulgarians rely on the social networks to receive news also shows the need for a more in-depth study of the level of media competencies of the country's audiences [20]. Thus media and digital literacy are among the prerequisites for media pluralism. The reason is that media and digital literacy guarantee access to more diverse sources of information.

Starting from the understanding that media literacy is a condition for universal access to information, for the development of critical thinking and for effective empowerment of citizens, the lack of media literacy policy is assessed as a risk to media pluralism.

4 Conclusion

In the hypermodern age, when technology is revolutionizing culture and it is "no longer in the representations, but in the objects, brands and technologies of the information society" [21], information and communication determine the parameters of the new "media" society. In order to sustain its proper functioning for the sake of deliberative communication combined efforts of all stakeholders (in the legal, regulatory, technological, economic, professional, academic, and social areas) are needed in all four domains. The findings in the review of the studied sources and databases and the conducted analysis of the media environment in Bulgaria (2000–2020) and the highlighted critical junctures can support outlining policies to enhance the perspectives for the media developments in the country.

Acknowledgements The research has been developed within the framework of the MEDIADELCOM research project of Horizon 2020 European Commission program.

References

1. MEDIADELCOM. https://www.mediadelcom.eu
2. Commission of the European Communities (9.10.2002) Regular report on the progress of the Republic of Bulgaria in the accession process. file:///C:/Documents%20and%20Settings/lili/My%20Documents/Downloads/rr_2002.pdf
3. Georgieva-Stankova N (2011) The "new" Bulgarian media—development trends and tendencies. Media regulation, ownership, control and the "invisible hand of the market". Trakia J Sci 9(3):191–203, Trakia University, Stara Zagora. https://www.academia.edu/3065709/_The_New_Bulgarian_Media_Development_Trends_and_Tendencies_Media_Regulation_Ownership_Control_and_The_Invisible_Hand_of_the_Market
4. Raycheva L (2013) The phenomenon of television—transformation and challenges. Tip-top Press, Sofia
5. Todorov P (2015) Deficits in the media labyrinth of the transition. https://www.unwe.bg/uploads/ResearchPapers/Research%20Papers_vol1_2015_No3_P%20Todorov.pdf
6. Smilova R, Smilov D, Ganev G (2011) Case study report. Does media policy promote media freedom and independence? The case of Bulgaria. Centre for Liberal Strategies (CLS). Academia.edu/8760427/Case_study_report_Does_media_policy_promote_media_freedom_and_independence_The_case_of_Bulgaria_Ruzha_Smilova_Daniel_Smilov_Georgy_Ganev_Centre_for_Liberal_Strategies_CLS_MEDIADEM?email_work_card=reading-history
7. Antonova V, Georgiev A (2013) Mapping digital media: Bulgaria. Report of the Open Society Foundation. file:///C:/Documents%20and%20Settings/lili/My%20Documents/Downloads/mapping-digital-media-bulgaria-en-20130805.pdf
8. Valkov I (2020) Without masks. Free Journalism. Annual Study on Freedom of Speech in Bulgaria. Association of European Journalists—Bulgaria (AEJ-Bulgaria). Statistical processing—Sociological Agency Alpha Research. https://aej-bulgaria.org/wp-content/uploads/2020/10/Jurnalisti-bez-maski-1.pdf
9. Television across Europe, More Channels, Less Independence (2008) Monitorung report. Open Society Institute, Budapest. https://www.opensocietyfoundations.org/uploads/3acad107-4566-48d6-bc17-1e6e98c00b41/1fullpublication_20080429_0.pdf
10. Raycheva L (2014) Mediaization of politics versus politicization of media in the situation of the election campaign. In: Krumov K, M. Kamenova, M. Radovic-Markovic (eds) Personality and society: the challenges of change. Bulgarian Academy of Sciences and Arts, Serbian Royal Academy of Sciences and Arts, European Center of Business, Education and Science, Sofia, pp 75–98
11. Reporters without Borders (2020) Press Freedom Index. https://rsf.org/en/world-press-freedom-index

12. European Commission (2021) E-government benchmark 2021. https://digital-strategy.ec.europa.eu/en/library/egovernment-benchmark-2021
13. European Commission (Eurostat) (2021) How digitalised Are EU's Enterprises? https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20211029-1
14. National Statistical Institute (2020) Culture. https://www.nsi.bg/en/content/3552/culture
15. European Commission (2021) Standard Eurobarometer: Report 92: media use in the European Union. https://op.europa.eu/en/publication-detail/-/publication/d2dbcf78-11e0-11ec-b4fe-01aa75ed71a1
16. Reuters Institute Digital News Report 2021 10th Edition (2021). https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf
17. Freedom House (2021) Freedom in the world 2021. Countries and territories. https://freedomhouse.org/countries/freedom-world/scores
18. Zankova B (2014) Governance, accountability and transparency of public service media in a contemporary mediatised world: the case of Bulgaria. In: Głowacki M, Jackson L (eds) Public media management for the twenty-first century: creativity, innovation and interaction. Routledge, New York and London, pp 125–142
19. Nieslen Admosphere (2021) Monthly Bulletin (04/2021), (05/2021), (06/2021). https://www.nielsen-admosphere.bg/products-and-services/tv-audience-measurement-in-bulgaria/audience-results/
20. Reuters Institute (2021) Digital news report. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021/dnr-executive-summary
21. Lash S (2004) Criticism of information. Kota Publishing House, Sofia, p 181

Towards an Adversary-Aware ML-Based Detector of Spam on Twitter Hashtags

Niddal Imam and Vassilios G. Vassilakis

Abstract After analysing messages posted by health-related spam campaigns in Twitter Arabic hashtags, we found that these campaigns use unique hijacked accounts (we call them *adversarial hijacked accounts*) as adversarial examples to fool deployed ML-based spam detectors. Existing ML-based models build a behaviour profile for each user to detect hijacked accounts. This approach is not applicable for detecting spam in Twitter hashtags since they are computationally expensive. Hence, we propose an adversary-aware ML-based detector, which includes a new designed feature (*avg_posts*) to improve the detection of spam tweets posted by the adversarial hijacked accounts at a tweet-level in trending hashtags. The proposed detector was designed considering three key points: robustness, adaptability, and interpretability. The new feature leverages accounts' temporal patterns (i.e., account age and number of posts). It is faster to compute compared to features discussed in the literature, and improves the accuracy of detecting the identified hijacked accounts by 73%.

Keywords Twitter spam detection · Adversarial examples · Evasion attack · Adversarial concept drift · Account hijacking · Trending hashtag

1 Introduction

The detection of Online Social Networks (OSNs) spam campaigns, which are accounts controlled by a malicious third party [8], has attracted researchers' attention not only because they irritate users, but also because these campaigns can be used to distribute more sophisticated security threats, such as malware or ransomware. Spam campaigns can create bots that are hard to be distinguished from regular users;

N. Imam (✉)
Alfaisal University, Riyadh, Saudi Arabia
e-mail: nimam@alfaisal.edu

V. G. Vassilakis
University of York, York, UK
e-mail: vasileios.vasilakis@york.ac.uk

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_32

these bots can easily generate a large number of spam tweets and spread misinformation to create a trending topic. In addition, spam campaigns evolve over time by adopting new techniques to evade detection [9, 10]. Spam campaign designers use different methods to fool the deployed spam detectors; for instance, using compromised (hijacked) accounts, creating fake accounts, or posting messages with empty content.

This paper proposes an approach for designing an adversary-aware ML-based detector of Twitter spam. After studying tweets posted by health-related campaigns on Twitter trending hashtags, we found that they use unique hijacked accounts as adversarial examples to fool spam detectors. Since most of the existing spam detectors, designed to detect hijacked accounts, need to analyse users' tweet history, which is not applicable approach in trending hashtags, we design a new feature *avg_posts* that can differentiate between legitimate user accounts and hijacked ones in our previous study [19]. The new feature leverages accounts' temporal patterns (i.e., account age and number of posts). Here, we developed an adversary-aware detector consisting of Multiple Classifiers System (MCS) for capturing different features of spam tweets with a Fuzzy Rule-based (FRB) classifier for aggregating the output of the classifiers and integrating the human-in-the-loop (HITL) approach. The developed adversary-aware detector was designed to be robust to identified adversarial examples by using the *avg_posts* feature, adaptable to evolving attacks (i.e., *adversarial drift*) and interpretable to enable experts (i.e., *security analysts*) to update the classifiers. The main contributions of this paper are as follows:

- An approach for designing an adversary-aware ML-based detector that is robust, adaptable, and interpretable is proposed.
- The robustness of the developed adversary-aware detector to the identified adversarial examples was evaluated and compared with state-of-the-art spam detectors.
- We demonstrate how the developed adversary-aware detector can handle the adversarial concept drift using a real-world dataset collected from Twitter.

1.1 Method

This section describes the methodology used for developing our adversary-aware ML-based detector of health-related spam in Twitter hashtags. It follows methods commonly used in the literature, but, in addition, the possible presence of adversaries was considered in each step. In our previous study [19] of health-related campaigns, we designed a new feature for detecting the identified hijacked accounts used by these campaigns. In this current study, we will use the designed feature *avg_posts* for developing an adversary-aware ML-based detector.

1.2 The Propose Adversary-Aware ML-Based Detector

Recent studies show that when ML-based models are used in security applications, they become vulnerable to various forms of adversarial attacks [1, 4, 5]. Thus, the developed adversary-aware detector was designed for an adversarial environment, in which the *adversarial drift* may occur because of adversaries' constant attempts to compromise cybersecurity systems. The developed adversary-aware detector of health-related spam, including the adversarial hijacked accounts, on Twitter is inspired by some of the models discussed in the related work section that utilize a *multiple classifiers systems (MCS)* [7, 20, 31, 34, 35]. Related studies focus on detecting hijacked accounts and tweets using MCS to capture different features (content-based or meta-based features) and to improve the detection accuracy. Unlike existing approaches, our adversary-aware detector consists of four classifiers: one meta feature-based and three textual feature-based classifiers. The detector considers four modalities of data: tweets' statistical features (e.g., account_age, status, avg_posts, etc.), tweets' textual content, tweets' description, and tweets' emojis content. Diversity is an important characteristic of MCS, as measuring the diversity helps prunes the classifiers [28]. According to [26], the best method to measure the diversity of a MCS is measuring the disagreement. Thus, the outputs of the four classifiers are feed into a *Fuzzy rule-based classifier (FRB)* for measuring the disagreement and making the final prediction. FRB systems consist of a set of IF...THEN rules that are transparent and interpretable by humans. FRB systems are widely used to deal with uncertainties or to process non-stationary streaming data [2, 11, 16]. Although the design of traditional FRB systems requires a number of handcrafting functions, assumptions and patterns to be selected, our detector utilizes an FRB system for integrating the outputs of the MCS in a way that does not require a large number of rules. The main reasons for using FRB are to detect possible adversarial drift that may occur as a result of adversarial attacks and to ensure the adaptability of the detector. Related studies use majority vote or softmax function for the final prediction of the MCS' outputs, which may not detect adversarial drift. We believe that when designing a spam detector for adversarial sittings, it is crucial to consider how the model will operate under a new adversarial attack. First, the input X is classified by the four classifiers; the output of these classifiers is either 0 (non-spam) or 1 (spam). The output of these classifiers is then examined by the FRB classifier, and the final decision of this classifier is the output Y. An overview of the developed adversary-aware detector is presented in Fig. 1. The following subsections will provide a more detailed description of the detector and its components.

2 Experimental Results

The experiments conducted in this section focus on initially choosing the best ML algorithms, and then training and testing the selected algorithms. As the developed

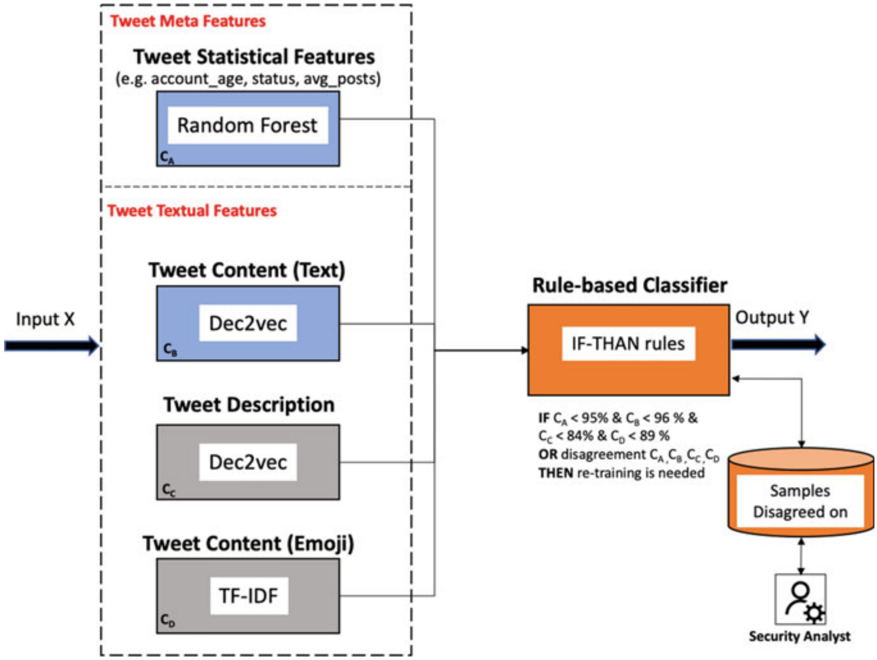


Fig. 1 An overall framework for spam detection

adversary-aware detector consists of four classifiers, separate experiments were performed to choose an ML algorithm for each classifier in the developed detector (see Fig. 1).

2.1 *Meta Feature-Based Classifier*

This classifier makes its predictions based on input statistical features, such as number of friends or followers; this classifier uses a total of 13 numerical features. The statistical features of tweets can help distinguish spam from non-spam. However, in the real world, these features may change in unpredictable ways over time, and several vectors may cause data distribution to drift over time. The meta feature-based classifier C_A focuses on detecting spam based on tweets' statistical features.

In our previous study [19] of the health-related campaigns, we found that they use a unique type of hijacked accounts. Thus, a new feature was designed $avg\_posts = \frac{status}{account\_age}$. Here, we focuses on evaluating the effectiveness of the *avg_posts* feature in improving the robustness of meta feature-based classifiers to the adversarial hijacked accounts with two goals. First, we seek to examine how well does the *avg_posts* feature improve the performance of different supervised and unsupervised classifiers in detecting the identified adversarial hijacked accounts. Second, we seek to compare our meta feature-based classifier with state-of-the-art spam classifiers.

Datasets We use three datasets in this part of experiments: Gilani-2017 [15] and cresci-rtbust-2019 [30] datasets, and the dataset collected from Twitter [18]. The latter consists of 2509 tweets that are grouped in two classes: 1990 non-spam tweets and 519 health-related spam tweets. The spam tweets includes 141 tweets that are posted by hijacked accounts. The three datasets were divided into training and testing datasets. Two versions of the training datasets were used; a training dataset that includes the adversarial examples and a training dataset that does not include adversarial examples. Hijacked accounts were considered as adversarial examples and used to evaluate the robustness of different ML-based models using different ML algorithms. Also, the benchmark datasets (Gilani-2017 and cresci-rtbust-2019) were used for evaluating the effectiveness of *avg_posts* by using different datasets.

Supervised approaches In this part of experiments, we compare the performance of six ML algorithms in detecting the adversarial hijacked accounts with and without the *avg_posts* feature. First, the ML algorithms were trained by using the training part of Twitter dataset that does not include the adversarial examples (i.e., cleaned dataset). Then, the algorithms were evaluated using the testing part of Twitter dataset. Results show that the overall prediction accuracy for most of the ML algorithms increases by at least 2% except for two ML algorithms when using the *avg_posts* feature. Based on the these results, we conclude that when the algorithms trained with dataset includes *avg_posts*, their performance on detecting the adversarial hijacked accounts increases.

Additionally, we conduct a preliminary experiment on *adversarial training*, by feeding the ML algorithms the training dataset that includes the adversarial examples. The goal is to evaluate the performance of the ML algorithms that trained on the adversarial training fashion, with and without the *avg_posts* feature. The results show that the overall detection accuracy of the ML algorithms is not affected when using *avg_posts*, yet it improves the detection accuracy in some algorithms.

Unsupervised approaches We compare the importance of *avg_posts* in the detection of adversarial hijacked accounts using models trained in an unsupervised manner as some recent studies used unsupervised approaches to detect hijacked accounts [20, 37] and bots [30]. Although supervised approaches can detect spam with high accuracy, their detection accuracy drooped on detecting never seen data (i.e., Zero-day attacks). We employ anomaly detection auto-encoders (AEs) [33] to evaluate the effectiveness of *avg_posts*. AE is a type of dimensionalty reduction and feature projection techniques (e.g., PCA, TICA). We used an AE as some related studies show that they outperform other diminsionality reduction techniques in detecting compromised accounts in OSNs [30, 37]. Similar to the above experiments, the collected dataset from Twitter was used for training and testing. Results show that there is a considerable improvement in the performance of the three unsupervised models when using *avg_posts*. The AUC of the dense-based AE improves by 12% when

Table 1 Comparison between our meta feature-based classifier and SOTA detectors

Algorithms	Recall
COMPA	0.19
Nauta	0.36
Botometer	0.71
Our classifier	0.73

using *avg_posts*, wheres 1 and 2% improvements were recorded for BiLSTM and LSTM, respectively.

Comparison Against Baselines Finally, we compare the detection accuracy of the adversarial hijacked accounts of our feature-based classifier against state-of-the-art spam detectors. The following three baselines were chosen: *COMPA* [12], *Nauta* [32], *Botometer* [3]. As most of related studies in detecting hijacked accounts use users' behavioural-based approach, which requires analysing users' tweet history, we build a new testing dataset that contains 52 users' accounts. The goal of these experiments is to show how these campaigns can fool hijacked accounts detectors that rely on accounts tweets history. We choose hijacked accounts that have old account age with very few tweets on their accounts (i.e., less than 10) and those that have only few spam tweets. These accounts are hard to be detected by users' behavioural-based detectors since their profiles do not have enough variations. After extracting the tweets of the 52 accounts, we manually evaluate them using COMPA's and Nauta's algorithms. The reason for the manual evaluation is that these algorithms cannot build a behaviour profile for accounts containing less than 10 tweets as stated in [12]. For evaluating the *Botometer*, we check each account in the dataset and record the results. If the score is higher than 3.5, the account is classified as spam. As the dataset contains only hijacked accounts, we compare the ability of the detectors to correctly find hijacked accounts (recall). The results in Table 1 show that our meta feature-based classifier outperforms the three detectors in detecting the adversarial hijacked accounts. For a fair comparison, we only used six features (no_followers, no_favourites, no_listed, status, account_age, and avg_posts) in this experiment. Although the *Botometer* detector achieves a result that is very close to ours, it classifies 34 out of 52 accounts with score and remarks that the "score might be inaccurate". The reason these accounts could not be classified accurately is that they have not been active for a long time and do not have enough variations.

2.2 *Content-Based Classifiers*

This subsection presents three parts: tweets' content, emoji and description classification. First, we compare the detection accuracy of the three text classifiers: Doc2vec [27], CapsuleNet [17], and BOW with TF-IDF[1]. Then, we examine the

[1] https://github.com/susanli2016/NLP-with-Python.

robustness of the three classifiers to a character-level type of attack. Followed by evaluating the detection accuracy of the chosen text classifiers using tweets' textual description. Finally, the results of emoji-based classifier are presented.

The first classifier was a Tweets' Content Classifier. We compare the detection accuracy of the three text classifiers: Doc2vec, CapsuleNet, and BOW with TF-IDF. The three classifiers were trained using the collected dataset from Twitter [18], which were split into training and testing datasets. The random forest algorithm was trained for the classification task. Results show that dec2vec achieves the best detection accuracy among the three classifiers. The rational explanation of this result is that dec2vec captures the meaning within embeddings [37].

An important step when designing an adversary-aware detector is to consider the robustness of the detector to adversarial examples. Since our analysis of the targeted campaigns reveals some adversarial activities that are carried out by these campaigns (e.g., adding repeated characters or misspelt spam words), we extend these types of character-level manipulation and create adversarial test dataset. First, we manipulated the top 30 frequent words in spam tweets by replacing some characters with visually similar symbols or numbers. For example, Maca 222 M@ca, Forever 222 F0rever. Then, we trained the classifiers using a clean version (i.e., does not contain adversarial examples) of the training dataset. Finally, we test the classifiers using the manipulated dataset. The results show that Dec2vec is the most robust classifier against the character-level attack. Based on these results, Dec2vec was chosen for the tweets' content classifier C_B.

Additionally, we a Tweets' Description Classifier since we found that the targeted campaigns use descriptions to mimic legitimate users' accounts. Our analysis shows that while a few, 12 out of 1990, non-spam tweets have empty description, 9 spam tweets have empty description. The Doc2vec with RF achieve 90% detection accuracy. The last classifier was a Tweets' Emoji Classifier. After striping emojis from tweets' content, and spliting the dataset into training and testing, we use TF-idf with RF for the classification C_D. The results show that our model can distinguish between the two classes with 98% detection accuracy.

2.3 Fuzzy Rule-Based Classifier

Here, a set of rules that depend on the outputs of the four classifiers are defined. The main reasons for using this classifier is to make sure that the detector can evolve over time in the face of emerging attacks. Specifically, the classifier was designed considering the adaptability and interpretability to handle possible adversarial drift that may occur as a result of adversarial activities [36]. To handle adversarial drift, two problems need to be considered: detecting possible adversarial drift and debugging /updating the detector. The proposed method for handling adversarial drift is a mix of active and passive approaches [28], in which we update the detector when the adversarial drift is detected (i.e., *active approach*) and when the classifiers disagree (i.e., *passive approach*). The methodology used for building this classifier was inspired by [35], which is one of the first studies that investigate the adversarial drift in streaming data. Based on the analysis of our dataset, we give the optional classifier C_C a higher score than C_D since we find that most of the tweets posted by accounts that has description. However, the sensitivity and weight of the classifiers can be updated if the disagreement between the classifiers increases. The FRB classifier will make its final decision based on the following three rules: (1) if both mandatory classifiers agree on an input class even if one or both optional disagree. (2) if one of the mandatory classifiers and both optional classifiers agree on an input class. (3) if one mandatory classifier agree with the optional classifier C_C. The output of the optional classifiers are considered only when the mandatory classifiers disagreed. These optional classifiers help overcoming the uncertainty and handling adversarial drift. Samples that the classifiers disagreed on, will be collected and used by the security analyst to update the classifiers.

Adversarial Drift Simulation To simulate the adversarial drift detection, we perform the following two steps:

Step (1) Detecting the adversarial drift:

1. Splitting the dataset into different chunks (D_1,..., D_n,...) where n is the number of chunks. The adversarial drift occurs between two points in time D_n^t and $D_{1_n}^{t+}$ where t is the time point. Each chunk contains a number of instances (i.e., tweets) $D_n^t = (x_1, \ldots, x_n, \ldots)$. if $P^{t(x,y)} \neq P^{t+1(x,y)}$, where $P^{t(x,y)}$ denotes the probability of data at a time point t, and y_n is the assigned to class of x_n.
2. Training our detector using clean dataset (i.e., not including adversarial examples)
3. Adding adversarial examples (hijacked accounts) to D_n^t with different percentages.
4. Evaluating our detector, which was trained on clean data, using testing datasets D_n^t.
5. The drift will be confirmed when the detection accuracy of the detector's classifiers dropped under the reference percentages.

In detail, we followed the methodology proposed in [35] for defining the reference percentages to which the predicted results of the classifiers were compared. The training dataset was used to find the expected accuracy for the classifiers. We uploaded the training dataset into WEKA and a tenfold cross-validation was chosen as a test option. After repeating this process ten times, the learned expected behaviours of the classifiers were used for adversarial drift detection. Classifiers' sensitivity to drift can be controlled by modifying the reference percentages. The reference percentages of our detector are as follows: C_A: 95%, C_B: 96%, C_C: 84%, C_D: 89%.

After choosing the reference percentages (i.e., accepted drift) of our classifiers, now we are simulating the adversarial drift on our dataset. The number of samples that are considered as an indicative of the adversarial drift is depending on the classifiers used. The experiment was preformed using the meta feature-based classifier C_A. In order for adversaries to inject an adversarial concept drift into data [22], they need to have knowledge about the nature of data. Thus, we will consider a scenario that starts with a *probing attack*, where an adversary manipulates a few samples (i.e.,

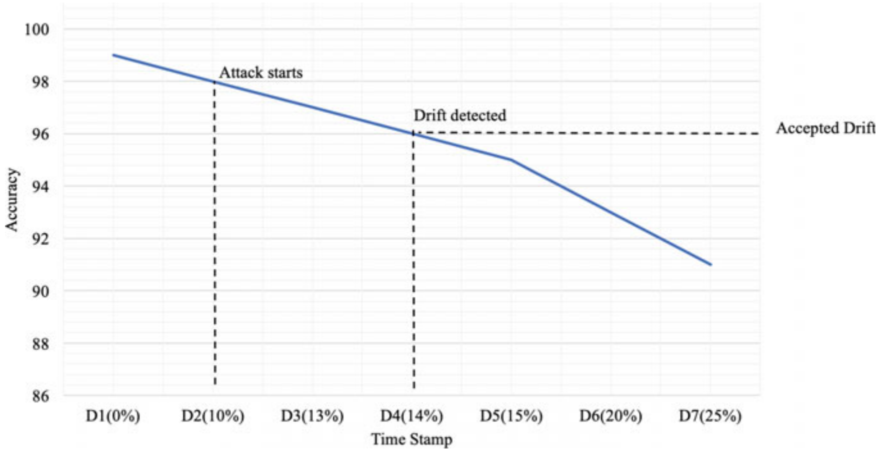


Fig. 2 The simulation of adversarial drift

tweets) and post them to learn from the deployed classifier's feedback. The next step is launching the adversarial attack, in which the adversary manipulates more samples to either evade detection or subvert the deployed classifier (i.e., Adversarial drift). To simulate the attack, we split the dataset into chunks ($D_1, \ldots, D_n, \ldots$). Each D_n^t denotes a set of samples that arrive at different time t. The training dataset consists of 400 non-spam and 100 spam tweets. The testing datasets consist of 350 tweets that include different percentages of adversarial examples (i.e., *adversarial hijacked accounts*). The manipulation percentages start from 10 to 25%. The simulation of the adversarial drift is presented in Fig. 2. The results show that the probing attack starts at D_2, where the manipulation percentage is 10%, and the drift is detected at D_4, where the manipulation percentage is 14%.

After simulating the detection of adversarial drift, the next step is to debug/update the classifiers. Different methods for collecting and labelling samples to update the deployed classifiers have proposed in the literature. Active learning focuses on choosing the most valuable data that need to be labeled, and has been widely used for solving this problem [21]. Several active learning methods are proposed to find the valuable samples, such as using uncertainty of a classifier [35], samples that best represent the concepts in distribution [13], or sliding windows [25]. Our proposed detector follows an active learning approach Query by Committee (QBC) [24] that finds the most valuable data to be used for updating the detector based on the disagreement between the classifiers. The QBC approach was first proposed for static active learning [14] and modified in [24] to be used for data stream. The labelling strategy we use is different from the one used by the adapted approach. We introduce our methodology for updating the detector in step 2.

Step (2) Updating and debugging the detector:

1. Samples that the classifiers (C_A, C_B, C_C, C_D) disagreed on $(x_1, \ldots, x_n, \ldots)$, where x_n is an input and n is the number of a sample that will be collected. These samples will be labelled by the FRB classifier using the fuzzy rules.
2. If the number of these samples reaches a certain level, they will be examined by the security analyst and used to evaluate the classifiers.
3. If the percentage accuracy of any of the classifiers dropped under the reference percentages, the collected samples will be used to update the detector.
4. Since the collected samples might not be sufficient for updating the classifiers, we will re-sample the collected samples by generating synthetic data [23].

The results in step 1 show that when we manipulate 10% of the data, the detection accuracy drops. Thus, based on this result, in which we use the statistical classifier C_A, if the number of samples that classifiers disagreed on is higher than 10% of the arrived chunks, the samples need to be checked by the security analyst. Updating the deployed detector requires a certain amount of human work although it may be costly and time-consuming [21]. Also, Ksieniewicz et al. [25] stated that for some practical tasks (e.g., medical diagnosis) humans need to verify labelled data. Hence, we integrate Human-in-the-loop approach in the process of updating the adversary-aware detector since the targeted type of drift occurs as a result of adversarial attack. Once the drift is confirmed by the security analysts, the collected samples will be used for debugging. There are different methods for debugging the classifiers, and in this research we consider retraining as the method of debugging. In some cases, retraining the classifiers may not enough and designing a new feature or using different ML algorithm is needed. If the collected samples were not sufficient for updating the classifiers, data re-sampling techniques that proved to be effective when dealing with few samples will be used. Synthetic Minority Oversampling Technique (SMOTE) [6], which is one of the most commonly technique used for oversampling, was chosen to generate new artificial samples by replicating pre-existing ones [29].

Additionally, we compare the accuracy and recall of two retraining methods used by the classifier C_A to handle the adversarial drift. We use the same setting for simulating the adversarial drift as in the previous experiment. After the drift is detected at $D_4(14\%)$, we used D_4 to retrain the classifier. We considered the classification accuracy and recall since the adversarial drift occurs as a result of manipulating the malicious samples only [35]. Then, we retrain the classifier using SMOTE from Imbalanced-Learn Library.[2] We over-sample D_4 and update the classifier. The results show that using SMOTE makes the recall more stable than updating the classifier using the detected adversarial drift's samples.

[2] https://github.com/scikit-learn-contrib/imbalanced-learn.

3 Conclusion

Motivated by the spread of untrustworthy healthcare advertisements in Arabic trending hashtags, we developed an adversary-aware detector of spam tweets posted by these campaigns. Excessive experiments on the collected datadets using the new *avg_posts* feature show that our adversary-aware detector outpreforms bots and hijacked accounts detectors. Additionally, the developed detector, which consists of MCS with a FRB classifier that integrates human-in-the-loop approach, was designed to be robust to the identified adversarial examples, adaptable to handle adversarial drift and interpretable to security analysts.

The aim of this study was to simulate the research community to focus on designing adversary-aware detection systems that are robust, adaptable, and interpretable. Although the analysis focused on spam campaigns in Arabic trending hashtags, as mentioned, the *avg_posts* feature can detect hijacked accounts regardless of the language used. Finally, achieving a high detection accuracy was not the main goal of this research, as the literature proves that, with enough data, it is not difficult to achieve high accuracy. Rather, our main focus was to develop adversary-aware spam detector keeping into accounts three key points: the robustness to the identified adversarial examples, adaptability and interpertability to handle adversarial drift (i.e., to ensure the detector can evolve over time).

References

1. Alabdulmohsin IM, Gao X, Zhang X (2014) Adding robustness to support vector machines against adversarial reverse engineering. In: Proceedings of the 23rd ACM international conference on conference on information and knowledge management, pp 231–240
2. Angelov PP, Gu X (2018) Deep rule-based classifier with human-level performance and characteristics. Inf Sci 463:196–213
3. Bessi A, Ferrara E (2016) Social bots distort the 2016 us presidential election online discussion. First Monday 21(11–7)
4. Biggio B, Corona I, Maiorca D, Nelson B, Šrndić N, Laskov P, Giacinto G, Roli F (2013) Evasion attacks against machine learning at test time. In: Joint European conference on machine learning and knowledge discovery in databases. Springer, pp 387–402
5. Biggio B, Fumera G, Roli F (2014) Security evaluation of pattern classifiers under attack. Knowl Data Eng 26(4):984–996
6. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority oversampling technique. J Artif Intell Res 16:321–357
7. Chu Z, Gianvecchio S, Wang H, Jajodia S (2010) Who is tweeting on twitter: human, bot, or cyborg? 6:10
8. Chu Z, Widjaja I, Wang H (2012) Detecting social spam campaigns on twitter. In: International conference on applied cryptography and network security. Springer, pp 455–472
9. Cresci S, Di Pietro R, Petrocchi M, Spognardi A, Tesconi M (2017) The paradigm-shift of social spambots: evidence, theories, and tools for the arms race. In: Proceedings of the 26th international conference on world wide web companion. International world wide web conferences steering committee, pp 963–972
10. Cresci S, Lillo F, Regoli D, Tardelli S, Tesconi M (2019) Cashtag piggybacking: uncovering spam and bot activity in stock microblogs on twitter. ACM Trans Web (TWEB) 13(2):1–27

11. Dou D, Jiang J, Wang Y, Zhang Y (2018) A rule-based classifier ensemble for fault diagnosis of rotating machinery. J Mech Sci Technol 32(6):2509–2515
12. Egele M, Stringhini G, Kruegel C, Vigna G (2013) Compa: detecting compromised accounts on social networks. In: NDSS
13. Ferreira RS, Zimbrão G, Alvim LG (2019) Amanda: semi-supervised density-based adaptive model for non-stationary data with extreme verification latency. Inf Sci 488:219–237
14. Freund Y, Seung HS, Shamir E, Tishby N (1997) Selective sampling using the query by committee algorithm. Mach Learn 28(2–3):133–168
15. Gilani Z, Farahbakhsh R, Tyson G, Wang L, Crowcroft J (2017) Of bots and humans (on twitter). In: Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining, pp 349–354
16. Gu X, Angelov PP (2020) Highly interpretable hierarchical deep rule-based classifier. Appl Soft Comput 106310
17. Hettiarachchi H, Ranasinghe T (2019) Emoji powered capsule network to detect type and target of offensive posts in social media. In: Proceedings of the international conference on recent advances in natural language processing (RANLP 2019), pp 474–480
18. IMAM N (2020) Health-related spam campaigns
19. Imam NH, Vassilakis VG, Kolovos D (2021) An empirical analysis of health-related campaigns on twitter arabic hashtags. Manuscript submitted for publication
20. Karimi H, VanDam C, Ye L, Tang J (2018) End-to-end compromised account detection. In: 2018 IEEE/ACM International conference on advances in social networks analysis and mining (ASONAM). IEEE, pp 314–321
21. Korycki Ł, Cano A, Krawczyk B (2019) Active learning with abstaining classifiers for imbalanced drifting data streams. In: 2019 IEEE international conference on big data (big data). IEEE, pp 2334–2343
22. Korycki Ł, Krawczyk B (2020) Adversarial concept drift detection under poisoning attacks for robust data stream mining. ArXiv preprint arXiv:2009.09497
23. Korycki, Krawczyk B (2020) Online oversampling for sparsely labeled imbalanced and nonstationary data streams
24. Krawczyk B, Woźniak M (2017) Online query by committee for active learning from drifting data streams. In 2017 international joint conference on neural networks (IJCNN). IEEE, pp 2120–2127
25. Ksieniewicz P, Woźniak M, Cyganek B, Kasprzak A, Walkowiak K (2019) Data stream classification using active learned neural networks. Neurocomputing 353:74–82
26. Kuncheva LI, Classifiers CP (2004) Methods and algorithms. Wiley, New York, NY
27. Le Q, Mikolov T (2014) Distributed representations of sentences and documents. In: International conference on machine learning, pp 1188–1196
28. Mahdi OA, Pardede E, Ali N, Cao J (2020) Fast reaction to sudden concept drift in the absence of class labels. Appl Sci 10(2):606
29. Maldonado S, López J, Vairetti C (2019) An alternative smote oversampling strategy for high-dimensional datasets. Appl Soft Comput 76:380–389
30. Mazza M, Cresci S, Avvenuti M, Quattrociocchi W, Tesconi M (2019) Rtbust: exploiting temporal patterns for botnet detection on twitter. In: Proceedings of the 10th ACM conference on web science, pp 183–192
31. Melis L, Song C, De Cristofaro E, Shmatikov V (2019) Exploiting unintended feature leakage in collaborative learning. In: 2019 IEEE symposium on security and privacy (SP). IEEE, pp 691–706
32. Nauta M (2016) Detecting hacked twitter accounts by examining behavioural change using twtter metadata. In: Proceedings of the 25th twente student conference on IT
33. Sakurada M, Yairi T (2014) Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2014 2nd workshop on machine learning for sensory data analysis, pp 4–11
34. Sculley D, Otey ME, Pohl M, Spitznagel B, Hainsworth J, Zhou Y (2011) Detecting adversarial advertisements in the wild. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp 274–282

35. Sethi TS, Kantardzic M (2018) Handling adversarial concept drift in streaming data. Expert Syst Appl 97:18–40
36. Sethi TS, Kantardzic M, Ryu JW (2018) Security theater: on the vulnerability of classifiers to exploratory attacks
37. VanDam C, Masrour F, Tan P-N, Wilson T (2019) You have been caute! early detection of compromised accounts on social media. In: Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining, pp 25–32

Higher Education Enterprise Resource Planning System Transformation of Supply Chain Management Processes

Oluwasegun Julius Aroba, Collence Takaingenhamo Chisita, Ndumiso Buthelezi, and Nompumelelo Mthethwa

Abstract The goal of this study was to outline the impact of the enterprise resource planning (ERP) system digital transformation of supply chain management (SCM) processes in higher education by using the desk research technique to gather information from other sources that we reviewed to build our study and identify gaps that were detailed in the discussions and results. This study concentrated on higher education, and observation was made that ERP systems do not fully cover all business operations, including supply chain management procedures such as price fixing, bid rigging, and collusion between employees and suppliers; yet the study satisfied all three research objectives by providing a recommended key methodology to enhance the ERP system of SCM integration in higher education.

Keywords ERP systems · Digital transformation · Supply chain management integration · Higher education

O. J. Aroba (✉) · C. T. Chisita
ICT and Society Research Group, Information Systems Department, Durban University of Technology, KwaZulu Natal, Durban 4000, South Africa
e-mail: Oluwaseguna@dut.ac.za; aroseglinks@yahoo.co.uk

C. T. Chisita
e-mail: collencec@dut.ac.za

O. J. Aroba
Honorary Research Associate, Department of Operations and Quality Management, Faculty of Management Science, Durban University of Technology, KwaZulu Natal, Durban 4001, South Africa

N. Buthelezi · N. Mthethwa
Audit and Taxation, Audit and Account Management Department, Durban University of Technology, Durban 4001, South Africa
e-mail: 21802607@dut4life.ac.za

N. Mthethwa
e-mail: 20720011@dut4life.ac.za

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_33

1 Introduction

In 1990s, Michael E. Porter introduced the “supply chain management” (SCM) to optimize the operations of supply chain management processes [1]. According to Nzama [2], “the improvement and the upgrading of the ERP program to advance the competitive advantage in the organizations might be presented with the new emerging risks due to the IT transformation”.

The supply chain digital transformation results in a more advanced automation and inter-system integration application, this implies that all machines and equipment in production are coordinated via the Internet and sensors to produce at the same time, and all necessary data is stored with the cloud system during this process [3].

The most common software used by institutions is *enterprise resource planning* (ERP), which requires a significant financial investment to set up and compared to other applications, and little research has been conducted on ERP systems in a university setting regarding keeping up with the constantly shifting expectations of the industry [4].

2 Problem Statement

The main purpose of these study is to evaluate the effects of transforming supply chain management processes using a higher education ERP system. These are following challenges that experienced by the higher education ERP systems and need be addressed by this study:

- Inability to adhere to the business requirements results poor evaluation and selection of ERP systems;
- Non-compliance with the legislative environment;
- Inadequate IT infrastructure; and
- Inadequate transfer of knowledge to embrace new technologies.

2.1 Research Objectives

The objectives of the study focus more to address the above challenges.

- Describe the pro-active planning provided by the ERP system to improve the supply chain management processes.
- Determine the effectiveness of the digital transformation in higher education supply chain management processes.
- Determine whether the ERP systems meet all the business requirements to promote the efficiency of business processes.

2.2 Research Questions

The study questions are prepared as follows, and the answers from these questions will depend on the research methodology;

- How does the ERP system enhance the supply chain management processes in higher education?
- What is the significance of digital transformation in higher education?
- What is the status of ERP system regarding compliance with all of the business requirements outlined in higher education policies, procedures, laws, and regulations?

3 Literature Review

ERP systems have been widely employed by major corporations worldwide, and they have recently supplanted management, financial, and administrative computer systems in higher education. ERP has played an important role in higher education's IT management, but it has been far from the core discipline of higher education [5].

Higher education institutions have failed to recognize the importance of the ERP system [6]. This is because there are very few successful implementations and adoptions of these applications, for example, in Australia, a recent study in 2020 found few institutions successfully implemented ERP system projects [7].

4 The Research Methodology

The study uses the qualitative research design to assess the impact of the higher education ERP systems in "digital" transformation of supply chain management processes. According to Rizkiana et al. [8], the qualitative method is a naturalistic research approach that employs a triangulation (combined) data collecting strategy with the researcher as the essential instrument.

4.1 Data Collection

The study uses the desktop research approach where we collected sources from the search engines to gather the existing journals or the work of other researchers to gain information that is relevant to our topic such as google search, DUT library search, google scholar, blogs, and any other online tracking tools.

4.2 Data Analysis

In this study, we reviewed the existing papers to identify gaps in the field, used the relevant information to build up our research, and achieved our research objectives. It involves discovering relevant patterns, pulling meaning from data, and establishing a logical chain of evidence—i.e., understanding how information is stored, processed, and interpreted.

4.3 Proposed Methodology to Improve ERP System of SCM Integration in Higher Education

Table 1 provides a detailed explanation of our prototype; ERP SCM integration (Fig. 1).

5 Business Process Integration

ERP is a system that can support several functions and merge them into a single database such as human resources, supply chain management, customer relationship management, finance, manufacturing functions, and warehouse management functions [8].

5.1 Evaluation of ERP System in Higher Education to Meet Business Processes

This study indicates that ERP is not simply an application but also a collection of other fundamental methodological issues.

Table 2 illustrates what is covered or/and not by the ERP solution for higher education [11].

5.2 How Does the ERP System Improves Supply Chain Management Processes?

ERP software can generate a bill of materials for all goods, track resources, and shipping paperwork and keep track of any last-minute modifications, this reduces "human mistake" and allows for speedier manufacturing, and ERP systems may help with packing procedures and quality inspections, as well as data management for customer shipments and invoicing [12] (Fig. 2).

Table 1 ERP SAP supply chain management integration [9]

University	ERP system development	Supply chain management
1. *Present a clear vision* Management needs to have a clear vison in terms of how to integrate their business processes into digital transformation in support of our higher education. It is critical to communicate a clear vision of what the digital transformation will do for supply chain now and the future	1. *Design principles*: Clear principles must be configured on the system with the aim to automate the business processes according to policies and producers implemented by management	1. *Student registration forms*: Higher education must insert parameters for students to fill out the forms online without any manual interference and must be easy to use
2. *Roles and capabilities*: Clear roles must be delegated to proficient users to take these new technologies and embrace it, to the advantage of higher education	2. *Standard user interface*: The higher education must provide specifications that are clear and suite the business processes of supply chain management	2. *Procurement processes*: Automated requests of goods and services, sourcing of quotations, creation of purchase orders, instruction on deliver terms and invoicing
3. *Change management*: Awareness must be conducted for staff to adapt to the new change by gradually introducing the new ERP system in procurement of goods and services. Also, an on-going training is necessary	3. *Feasibility study*: The study must be conducted to understand the compactible modules to be used to improve the processes of supply chain management	3. *Market user interface*: The higher education must link their systems with the market related prices for demand management and supply of goods and services. This section will avoid price fixing and bid rigging as the system with reject over pricing and projects exceed the final budget
4. *ICT Steering Committee*: Higher education needs to have a strong ICT committee to deal and assess every system procured by the organization with aim of value for money and return on investment. The committee must lead in the implementation of projects	4. *Integrated processes*: Seamless integration is required to ensure that the system is compactible to work with other systems within the organization	4. *Online payment*: All purchase orders must be paid after comparing the invoice against the order in the system, and the system must provide a proof of all goods delivered. Payment must be made as per two authentication signatories

5.3 Digital Transformation in Supply Chain Management Processes

Companies with greater end-to-end visibility into the complexity of their supply chains and logistics operations, as well as digitally transformative processes and systems, provide accurate, timely, and incomplete access and transparency to events and data for transaction, content, and related supply chain information, both within

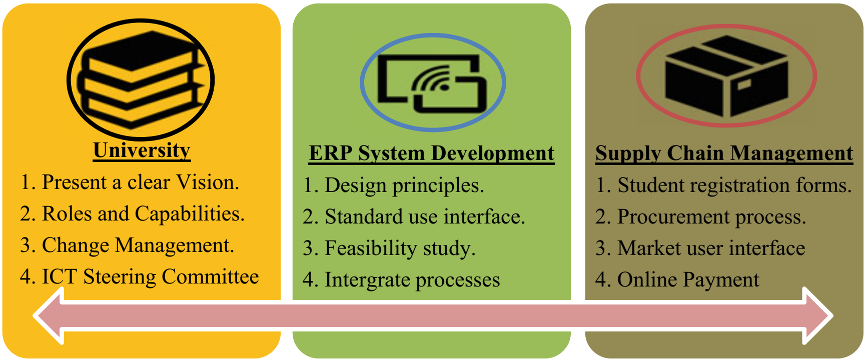


Fig. 1 Proposed methodology to improve ERP System of SCM Integration in higher education [10]

and across organizations, and support the effective planning and execution of supply chain operations [13].

Supply chain management in "digital" transformation is more than just deploying new technology; it is about leveraging new technologies to radically change how your company runs and provides value to its customers, potential benefits of a fully realized digital supply chain include savings across the board, such as reduced time, resources, money, and environmental footprint [14].

5.4 *Advantages and Disadvantages of Using ERP Solutions in SCM Processes*

The advantages of employing ERP solutions in SCM processes to improve the functions implemented by the institution of high education as listed below:

- ERP functional modules all serve various business functions, but the most beneficial system feature for supply chains is unquestionably simple integration.
- Provides business with automation of purchasing product or services with a strong competitive advantage.
- Embrace the new technologies that are coming to the market and meet customer needs.

Table 2 illustrates solutions covered or/and not covered by the ERP systems

Seq.	Item	Description	Tasks covered by ERP	Tasks NOT covered by ERP	Impact factor%
1	Strategy	Vision, mission, strategic, objective, goals	No	Yes	15
2	Business	Corporate policies, operating model, business process, bylaws important	No	Yes	10
3	Data structure	Data models: conceptual, logical, and physical	Yes	Yes	10
4	ERP application	Application software pool of data and knowledge	Yes	No	35
5	Workforce	Employee assessment	No	Yes	5
6	Facilities	IT infrastructure assessment			
		Clients Network Storage Application Data Security Change Project management IT administration	No No No Yes Yes No No No No	Yes Yes Yes No No Yes Yes Yes Yes	
7	Services	SLA/SLM "service level management" to secure the corporate investments	No	Yes	10
8	Training	On-going training plan for the whole staff in different levels	No	Yes	N/A

Fig. 2 Image illustrates the ERP system enhancing the SCM processes [12]

Disadvantages of using the ERP solution in the SCM processes are listed below:

- ERP solution failure to be compatible with the business processes and complicate the functions of the organizations.
- Businesses need to adapt on the EPR solution which results the EPR system not in compliance with the company's policies, procedures, laws, and regulations.
- ERP systems are not fully utilized due to the fact that they become more complex to the users.
- Business do not have the owning rights toward their information archived to ERP solutions.

6 Results and Discussion

The findings of our study are based on the research questions which will inform our recommendation and conclusion. There have been several good studies that address strategic planning concerns, when evaluating a stymied or failed ERP installation and determining the factors that caused it to fail. Frequently, the university administration will determine that the program does not work or is too complicated to apply in their specific setting [15].

Table 2 indicates that the organizations purchase ERP solutions for the sake of buying it but not utilizing for its full capacity. The current corporate ERP systems offer a distinct set of features that range dramatically from the academic functionality required by higher education institutions, which are not compactible from the business processes employed by higher education as a result ERP for higher education does not specially address academic functions; therefore, ERP for higher education should begin with the organization structure, which includes strategy/policy, data flow, business process structure, and academic functions.

ERP system improvements in supply chain management lack features that prevents improper behaviors' such as price rigging, specification fixing by end users, and collusion between suppliers and administrators for personal gain or fraud, resulting in organizations losing a lot of money due to poor supply chain management, and the ERP system does not cover all of that.

Another issue of digitalizing the supply chain process is poor of technical support within the organization as the EPR system applications are outsourced and results the higher education incur excessive expenditure due to lack of capacity skills. In this study, we also discovered that almost all organizations do not have full ownership of their data communicated in the ERP system because they do not have all of the system's rights, saved by the service providers in case where the organizations failed to pay their monthly subscriptions.

7 Recommendations and Conclusions

Based on our conclusion for this study, we should not underestimate the significance and quick expansion of digital revolution in higher education. In our literature review, it emphasized that the ERP systems have been employed by the biggest corporations worldwide, but the higher education has failed to recognize its importance; therefore, this study developed the methodology which present a clear strategy that need to be employed to improve ERP systems of SCM integration in higher education.

The methodology applied addressed the findings identified on the study, which shows that the ERP systems play a huge role in the development of transforming in supply chain management systems, and even though the implementation of the ERP systems has encountered some challenges such as features that are not compatible to the business processes and issue of non-compliance with the company's policies, laws, and regulations. Universities should define clear vision and goals in deploying ERP systems to achieve seamless integration and improve supplier chain management processes.

References

1. Qui M (2022) Research on book purchase supply chain management of university library. Int J Organ Innov 14(4):390–398
2. Nzama L (2021) Enterprise resource planning (ERP) upgrades in South Africa higher education institutions. J Glob Bus Technol 17(1):98–109
3. Özkanlısoy Ö, Akkartal E (2021) Digital transformation in supply chains: current applications, contributions and challenges. Bus Manage Stud Int J 9(1):32–55
4. Kumar M, Garg A, Kumar A (2021) Critical factors of post implementation of ERP in higher education system survey review. IOP Conf Ser Mater Sci Eng 1149(1): 3–9
5. Mohamed Hashim MA, Tlemsani I, Matthews R (2022) Higher education strategy in digital transformation. Nat Publ Health Emerg Collect 27(3):3171–3195
6. Ramchandran N, Thangamani G (2020) Factors for implementation of ERP in higher education—a literature review. In: Seventeenth AIMS international conference on management, pp 337–339
7. Abugabuh A (2020) ERP systems in higher education: a literature review and Implications. Signif Contrib 15(1):395–390
8. Rizkiana A, Ritchi H, Adrianto Z (2021) Critical success enterprise resource planning (EPR) implementation in higher education. J Account Audit Bus 4(1):56–65

9. Aroba OJ, Chinsamy KK, Makwakwa TG (2023) An ERP implementation case study in the South African retail sector. In hybrid intelligent systems: 22nd international conference on Hybrid Intelligent Systems (HIS 2022), December 13–15, 2022. Cham: Springer Nature Switzerlandsmart, pp 948–958. https://doi.org/10.1007/978-3-031-27409-1_87
10. Aroba OJ, Mnguni SB (2023) An Enterprise Resource Planning (ERP) SAP implementation case study in South Africa small medium enterprise sectors. In: Motahhir S, Bossoufi B (eds) Digital technologies and applications. ICDTA 2023. Lecture notes in networks and systems, vol 668. Springer, Cham, pp 348–354. https://doi.org/10.1007/978-3-031-29857-8_35
11. Noaman Y, Ahmed F (2020) ERP functionalities in higher education. Procedia Comput Sci 65:392–395
12. Omni (2020) Why consider ERP supply chain management. Available at: https://www.omniaccounts.co.za/why-consider-erp-supply-chain-management/. Accessed on 22 Sept 2022
13. Özkanlısoy Ö, Akkartal E (2021) Digital transformation in supply chains: current applications, contributions and challenges. Bus Manage Stud Int J 9(1):35–55
14. Singh S (2022) Supply chain digital transformation: how and why it matters in an organization. Available: https://appinventiv.com/blog/digital-transformation-in-supply-chain-management/. Accessed on 28 Sept 2022
15. Asprion PM, Scheinder B, Crimberg F (2022) ERP systems towards digital transformation: abstract. ERP Syst 1(1):45–109

Reduced Complexity Iterative LDPC Decoding Technique for Weak Atmospheric Turbulence Optical Communication Link

Albashir A. Youssef

Abstract In recent years, much research has been done on free-space optical communication (FSO). The unregulated spectrum, low implementation costs, and robust security of FSO systems are some of the reasons for this consideration. However, the fundamental limitation with FSO links is atmospheric turbulence (AT). Random phenomena are the best characteristic by turbulence of atmosphere caused by changes in the air's refractive index over time. Channel coding is one of the possible solutions for mitigating such FSO channel impairments as the low-density parity check (LDPC) code. In this article, the implementation efficient reliability ratio weighted bit flipping (IERRWBF), the modified IERRWBF (MIERRWBF), and weighted bit flipping (WBF) techniques are compared and evaluated against FSO atmospheric turbulence channels. The results show an impressive improvement of the coded FSO system by employing the MIERRWBF technique compared to the uncoded one from the point of all considered comparison parameters.

Keywords FSO · WBF · LDPC

1 Introduction

Optical carriers make it possible to explore new opportunities in wireless communications still not explored yet. Integrating electromagnetic waves-based wireless communication systems with optical carriers with. It will have a significant impose on enabling supporting future-generation heterogeneous wireless communications. It will support an expanded range of applications and services.

FSO systems employment still imposed various limitations. These significant issues are turbulence due to the atmosphere, attenuation impacts due to weather, and geometric losses. The laser beam scintillation is due to atmospheric turbulence caused by the refractive index differences in atmosphere. The pressure, temperature,

A. A. Youssef (✉)
Arab Academy for Science, Technology and Maritime Transport, Cairo, Egypt
e-mail: albashir.adel@aast.edu
URL: https://www.aast.edu

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_34

and wind variations are the leading causes of these differences [14]. Modeling of turbulence of atmosphere is performed by statistical models that suit the experimental results. Taking into account, model of log-normal [17] for weak turbulence of atmosphere regime. The turbulence due to atmosphere and weather conditions attenuation due to dust, rain, fog, snow, and haze causing fading to lead to a considerable influence on performance of the optical system [1, 4].

Various modulation techniques are maintained in FSO systems for minimizing atmospheric turbulence consequences according to energy or spectral efficiencies and non-coherent or coherent detection. The considerable typical techniques are on-off keying (OOK) [10], pulse position modulation (PPM) [24], pulse width modulation (PWM) [5], multiple PPM (MPPM) [23], binary phase-shift keying (BPSK) [15], space shift keying [12], and digital pulse interval modulation (DPIM) [9]. A new multipoint-to-multipoint signal-space diversity (SSD) cooperative FSO scheme is delineated and investigated under log-normal and gamma–gamma channel models for various users utilizing variant levels of modulation [12]. In [21], various modulation scheme performances (combined or not combined with space diversity reception technique (SDRT)) employed in systems of FSO are investigated.

The error control coding techniques are the most promising mitigating processes for the atmospheric turbulences of FSO channels. In [18], polar codes are analyzed for their performance for selecting the excellent rate of the code needed to reach a 10^{-9} bit error rate at weak atmospheric turbulence. Also, in [7], polar codes are introduced and compared to low-density parity check (LDPC) codes. According to maintained simulation results, LDPC codes achieve lower BER than polar codes. Authors in [11] assess uncoded and coded FSO communication system performance as they utilize Bose Chaudhuri Hocquenghem (BCH) and LDPC codes. The study performed in [2] shows that using a dynamically adjusted log-likelihood ratio (LLR) technique is characterized as a soft decision technique. Soft decision techniques are known for their impressive coding gain performance, while their complexity is immense.

As most recent published works concentrate on soft decision LDPC decoding techniques as the best candidate for FSO channels. In [25], more enhanced hard decision techniques are proposed for enhancing FSO channels performance and perform close to soft decision ones. The techniques introduced in this study are weighed bit filliping (WBF) and implementation efficient reliability ratio weighed bit filliping (IERRWBF). The latter one performed better than WBF against weak and moderate atmospheric turbulence.

Further improvement is required for the atmospheric turbulence channels, so a more enhanced hard LDPC decoding technique which is proposed in this paper is MIERRWBF. It lowers the complexity of the whole FSO coding system.

However, to the author's knowledge, no recent attempts in recently published works concern recently proposed LDPC decoding techniques such as MIERRWBF for enhancing the FSO atmospheric turbulent channels. Also, the Monte Carlo simulation results for the recently proposed technique results in impressive enhancement in the coded FSO communication system compared with other techniques from complexity.

The organization of the paper is as following: Sect. 2 shows model of the proposed system. Section 3, the FSO channel model is presented. In Sect. 4, techniques of LDPC encoding and decoding are illustrated. Simulation results are shown in Sect. 5. Conclusion is illustrated in Sect. 6.

2 Proposed System Model of FSO Communication

As presented in Fig. 1, the system model proposed for FSO communication illustrates that the binary source data will be LDPC coded and mapped by the on-off keying technique. The resultant electrical circuit will be transformed into an optical beam using the photodiode on the transmitter side. The optical beam transmitted will be exposed to weather attenuation causing turbulence of atmosphere and path losses. The analytical expression of electrical signal r(t) received is

$$r(t) = y(t)\,\eta\, I + n(t), \tag{1}$$

$$I = \beta I_o h, \tag{2}$$

So, $y(t)$ is the electrical signal transmitted, η represents responsively of detector, $n(t)$ is the additive white Gaussian noise (AWGN) which has σ_n^2 (variance) equal to $N_o/2$ and zero mean, and I is the signal's received intensity which is illustrated by [3].

3 Weak Atmospheric Turbulence FSO Channel Model

A mathematical channel model has been proposed to specify turbulence of atmosphere for weak case. The channel model is delighnated by its probability density function (pdf) which is shown in Eq. 3 [19].

$$f_{h_i}(h_i) = \frac{1}{h_i\sqrt{8\pi\sigma^2}} \exp\left(-\frac{(\ln(h_i) - 2\mu)^2}{8\sigma^2}\right), \tag{3}$$

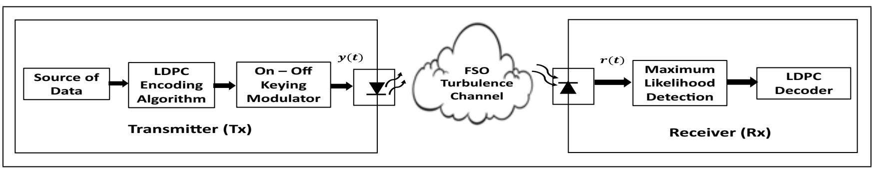


Fig. 1 Proposed FSO system model

as the coefficient of channel is $h_i = \exp(2Z)$ characterized by Z being an independent and identically distributed (i.i.d.) Gaussian random variable (RV) with mean μ, standard deviation σ, and variance σ^2. For ensuring that the fading channel not causes attenuate or amplifies the average power, the fading coefficients are normalized as $\mathbf{E}[{h_i}^2] = e^{2(\mu+\sigma^2)} = 1$.

4 Techniques of LDPC Encoding and Decoding

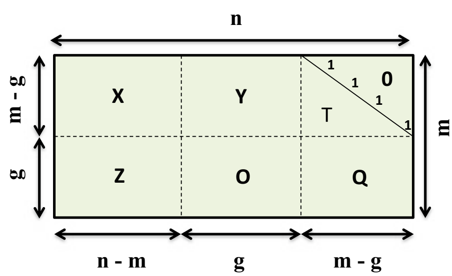
4.1 Encoding of LDPC

The LDPC codes construction mainly relies on a parity check matrix characterized by sparseness features. So, an efficient encoding procedure proposed in [20] is maintained using the parity check matrix, which will be applied as an alternative for converting the matrix of parity check into a generator matrix, which does not affect its feature of sparseness associated with **H** matrix. It will causing extra complexity in encoding [20]. The outcome of this process is an exact matrix shape with lower triangular as appears in Fig. 2.

4.2 LDPC Decoding Techniques

There are three types of LDPC decoding techniques. First category concerns techniques with hard decision. The second category is the soft decision techniques characterized by immense complexity with impressive BER. Finally, the third category is the hybrid decoding techniques which compromise between lower complexity of hard ones and outstanding BER performance of soft ones.

Fig. 2 Lower triangular form



4.3 LDPC Hard Decision Decoding Techniques

Weighted Bit Flipping (WBF) The WBF technique is proposed in [26]. It seeks to improve the error correction ability of the decoding technique termed by BF proposed in [8] by retaining the good correct ability for the data symbols in its decisions of decoding. Therefore, the additional complexity of decoding is obligatory to reach enhancement in performance.

The WBF decoding starts by recognizing considerably inaccurate variable nodes linked to every check node. This step is defined by following equation:

$$\mid y_{n_{\min}} \mid = \{\min \mid y_n \mid : n \in \mathcal{N}(m)\} \tag{4}$$

as the $n_{\min}$ represents the index of the lesser soft value of variable nodes linked to the check node m.

The minimum absolute component in the sequence received is calculated which is $\mid y_n \mid$, characterizing the reliability calculation of message received [26]. Since the binary counterpart b_n to y_n as its soft value with formidable reliability $\mid y_n \mid$. It is binary digit b_n which is leveled up. The determination of error-term E_n for each variable node is expressed by

$$E_n = \sum_{m \in \mathcal{M}(n)} (2s_m - 1) \mid y_{n_{\min}} \mid \tag{5}$$

as the syndrome associated bit s_m linked to m check node. The E_n represents the weight checksum which is connected to the n code bit. The procedure of the WBF technique is thoroughly explained in Table 1.

Implementation Efficient Reliability Ratio Weighted Bit Flipping (IERRWBF) It is observed that the RRWBF proposed technique by [6] consumes lots of operations, so a vital modification is performed to minmize the RRWBF technique complexity and keep the improvement in BER over the WBF technique. So a lower complexity calculation term is proposed in [16]. This term target lessening the decoding time consumed in the RRWBF technique is proposed by using T_m instead of the reliability ratio factor:

Table 1 Steps of WBF decoding

Step 1	If $\mathbf{s} = \mathbf{z}\,\mathbf{H}^T$ results in all are zero, halt the decoding
Step 2	Calculate E_n according to (5), for $1 \leq n \leq N$
Step 3	Distinguish the bit position n where E_n is the largest
Step 4	Flip the hard decision of y_n represented by z_n
Step 5	Steps 1–4 will be repeated as far as all the parity check equations are satisfied, or a pre-set maximum number of iterations is achieved

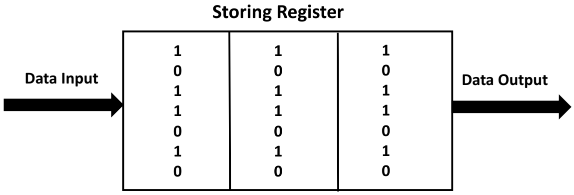


Fig. 3 Three entry Register.eps

$$T_m = \sum_{n \in \mathcal{N}(m)} \mid y_n \mid \tag{6}$$

also calculation of E_n as follows:

$$E_n = \frac{1}{\mid y_n \mid} \sum_{m \in \mathcal{M}(n)} (2s_m - 1)T_m \tag{7}$$

Modified Implementation Efficient Reliability Ratio Weighted Bit Flipping (MIERRWBF) The main shortcoming of the latter iterative decoders is the expenditure of the extended time in the decoding process, especially at the variable node and check node steps without any additional enhancement in BER [27]. The decoding technique IERRWBf proposed in [16] suffers from this primary concern.

The technique termed MIERRWBF proposed In [27] added a decision step to figure out situations illustrated in the last section and restrict the loop of iterations by selecting either proceeding with decoding or termination loop of iteration results in enhanced decoded word. The phenomena of oscillation examined in the last paragraph are demonstrated in Fig. 3 for additional clarification.

4.4 Soft Decision LDPC Decoding Techniques

Min-Sum Technique Techniques with the soft decision are extracted from the proposed technique in [8] termed by belief propagation (BP) technique. These techniques are distinguished by complexity of $O(2M\rho + 4N\gamma)$ for each decoding iteration [13]. Decoding techniques with lessening complexity are extracted from the BP technique, which is the min-sum technique proposed in [13]. The procedure of min-sum decoding is illustrated in Table 2.

5 Simulation Results

Results of simulation are shown in this section to validate the derived analysis in this paper and prove improvement due to operating recently proposed decoding technique termed by MIERRWBF. In all conducted analyzes, the following parameters are

Table 2 Steps of min-sum decoding

Step 1	If $\mathbf{s} = \mathbf{z}\,\mathbf{H}^T$ results in all are zero, decoding will be halted
Step 2	Initialization $F_n = \frac{4}{N_o}\, y_n$ where N_o is spectral noise power density
Step 3	Horizontal step computed by: $L_{mn} \approx \prod_{n' \in \mathcal{N}(m)\bar{n}} sgn(z_{mn}).\min_{n' \in \mathcal{N}(m)\bar{n}} \mid z_{mn'} \mid$
Step 4	Vertical step computed by : $z_{mn} = F_n + \sum_{m' \in \mathcal{M}(n)\bar{m}} L_{m'n}$
Step 5	For binary format conversion: $z_n = F_n + \sum_{m' \in \mathcal{M}(n)} L_{m'n}$
Step 6	Syndrome $\mathbf{s} = \mathbf{z}\,\mathbf{H}^T = 0$ stop decoding

Table 3 System configuration [22]

Parameter	Symbol	Value
Wavelength	λ	1550 nm
Receiver diameter	D_{R}	0.2 m
Transmitter diameter	D_{T}	0.2 m
Divergence angle	θ_{T}	2 mrad
Separation between the source and the destination	L	1 km
Coefficient of attenuation	α	0.43 dB/km
Jitter standard deviation	σ_{s}	0.3 m
Beam waist	w_{z}	2 m
Pointing error parameter	ξ	3.3377
Refractive index constant (weak atmospheric turbulence)	C_n^2	$0.5 \times 10^{-14}\ \mathrm{m}^{-2/3}$

considered in the maintained simulation in this paper. The BER targeted for FSO channels is 10^{-6}, $\alpha = 0.43$ dB/km, $\lambda = 1550$ nm, $\ell = 1000$ m, for conditions of clear weather and strong sunlight conditions. In simulation results, for each $E_b/N_o 10^7$ bits are transmitted. The utilized parameters in the results of simulation are demonstrated in Table 3.

In Fig. 4, BER is compared for recently proposed techniques and other published techniques concerning enhancing the BER of FSO weak atmospheric turbulence channel. As delineated in Fig. 4, the MIERRWBF achieves the same BER levels as the IERRWBF technique at all maintained $E_{\mathrm{b}}/N_{\mathrm{o}}$s. Besides, it gets close to the soft decision technique termed a min-sum technique characterized by superior BER performance.

Another factor for evaluating the LDPC decoding techniques over FSO atmospheric turbulent channels is the average iterations number consumed by each decoder. According to Fig. 5, the average iterations number versus $E_{\mathrm{b}}/N_{\mathrm{o}}$ is interpreted for weak atmospheric turbulence channel. It is noticed that the required average iterations number belonging to MIERRWBF techniques reached the bottom of the number of iterations compared to other techniques under study, especially at the $E_{\mathrm{b}}/N_{\mathrm{o}}$s from 8 to 10 dB.

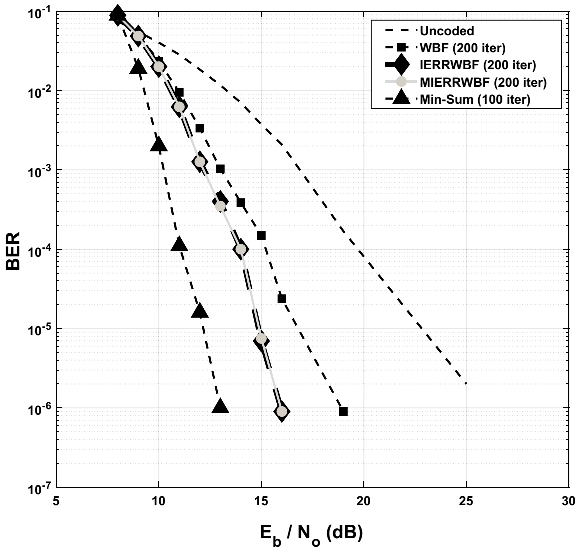


Fig. 4 BER comparison between LDPC decoding techniques for proposed system

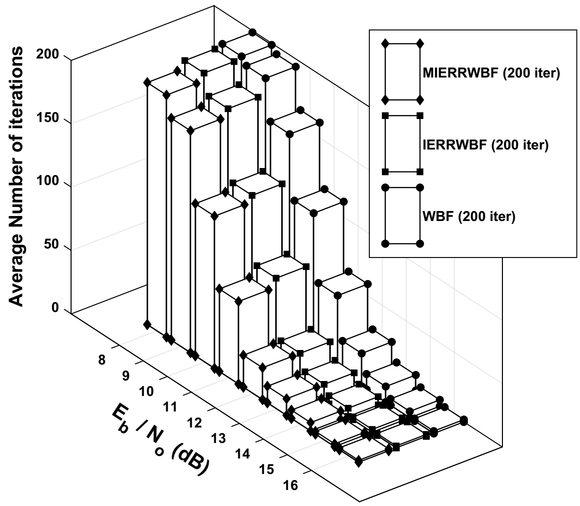


Fig. 5 Average number of iterations comparison between LDPC decoding techniques for proposed system

The decoding computation time for all maintained techniques is compared along weak atmospheric turbulence channel. Fig. 6 shows the comparison between all maintained LDPC decoders from the point of decoding computation time over the weak atmospheric turbulence channel. In Fig. 6, it is observed that the lowest level of decoding computation time belongs to MIERRWBF compared to other techniques all over the E_b/N_os, saving the wasted computation time at other techniques discussed in the later paragraphs due to its successful stopping criterion illustrated in the later sections.

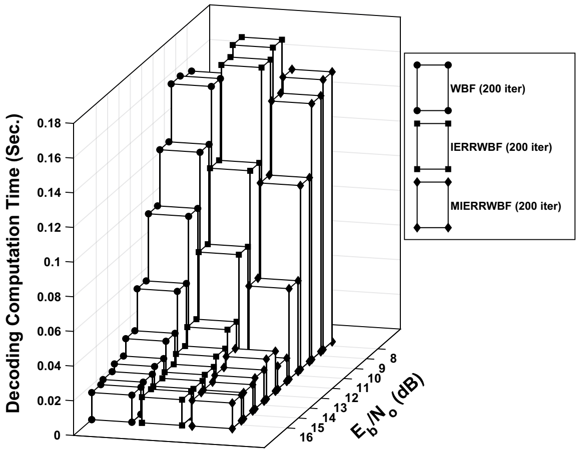


Fig. 6 Decoding computation time comparison between LDPC decoding techniques for proposed system

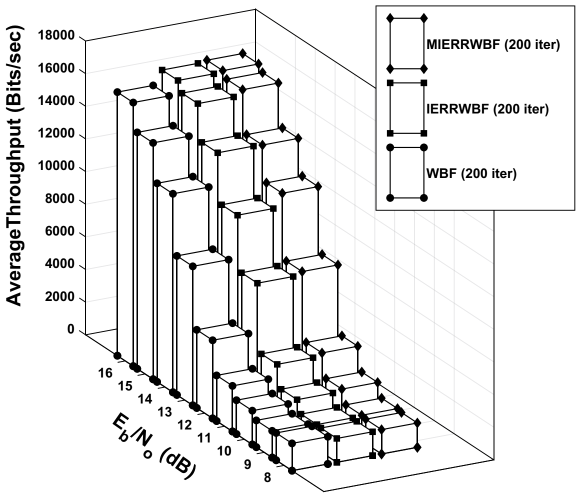


Fig. 7 Average throughput comparison between LDPC decoding techniques for proposed system

The resultant average throughput is a crucial parameter in evaluating LDPC decoding techniques against FSO atmospheric turbulence channels. The average throughput comparison for LDPC decoding techniques under study at weak atmospheric turbulence channel is presented in Fig. 7. At $E_b/N_o = 11$ to 13 dB, the MIERRWBF technique reached the highest average throughput over all maintained technique. This variation is due to the variant performance of the weak atmospheric turbulence channel. The average throughput at the same turbulent channel all techniques under study saturated by the same average throughput value exactly from $E_b/N_o = 8$ to 10 dB. The IERRWBF technique maintained the lowest average throughput at most of E_b/N_os.

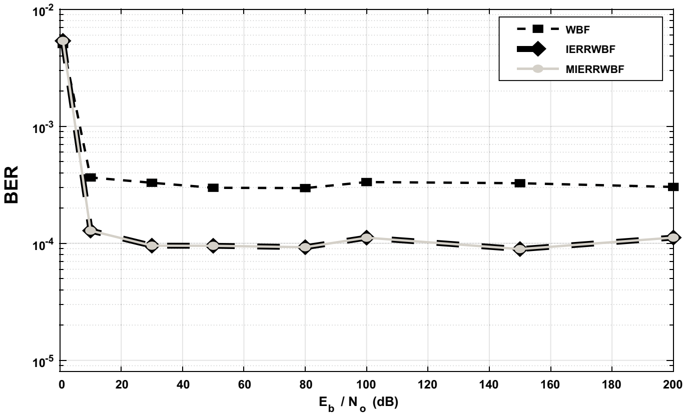


Fig. 8 Convergence comparison between LDPC decoding techniques for proposed system

Convergence is a vital parameter that concerns iterative decoding techniques evaluation. It is noticed from Fig. 8 that MIERRWBF achieved the fastest convergence at channel of weak atmospheric FSO turbulence . The BER of the weak atmospheric had the lowest converged BER, approximately 4×10^{-5}.

6 Conclusion

This paper evaluates various LDPC decoding techniques in a weak atmospheric turbulence channel of FSO communication systems. LDPC decoding techniques have three decision categories: hard, soft, and hybrid; all are considered in this evaluation to select the best suited FSO atmospheric turbulence channel. The evaluated performance considers crucial parameters for comparing LDPC decoders performance metrics are based on BER, the average iterations number, convergence, decoding computation time, and average throughput. The weak atmospheric turbulence channel model is considered in this evaluation. The MIERRWBF technique maintains impressive performance against all evaluation parameters considered in this work.

References

1. Bayaki E, Michalopoulos DS, Schober R (2012) EDFA-based all-optical relaying in free-space optical systems. IEEE Trans Commun 60(12):3797–3807
2. Cao P, Rao Q, Yang J, Liu X (2021) LDPC code with dynamically adjusted LLR under FSO turbulence channel. J Phys Conf Ser 1920:012023
3. Mohammad Taghi Dabiri and Seyed Mohammad Sajad Sadough (2018) Performance analysis of all-optical amplify and forward relaying over log-normal FSO channels. J Opt Commun Netw 10(2):79–89

4. Datsikas Christos K, Peppas Kostas P, Sagias Nikos C, Tombras George S (2010) Serial free-space optical relaying communications over gamma-gamma atmospheric turbulence channels. J Opt Commun Netw 2(8):576–586
5. Fan Y, Green RJ (2007) Comparison of pulse position modulation and pulse width modulation for application in optical communications. SPIE Opt Eng 46(6)
6. Feng Guo, Hanzo Lajos (2004) Reliability ratio based weighted bit-flipping decoding for low-density parity-check codes. Electron Lett 40(21):1356–1358
7. Fujia S, Okamoto E, Takenaka H, Kunimori H, Endo H, Fujiwara M, Shimizu R, Sasaki M, Toyoshima M (2021) Performance analysis of polar-code transmission experiments over 7.8-km terrestrial free-space optical link using channel equalization. In: International conference on space optics-ICSO 2020, vol 11852, pp 2301–2310. SPIE
8. Gallager Robert G (1962) Low-density parity-check codes. IRE Trans Inform Theor 8(1):21–28
9. Ghassemlooy Z, Hayes AR, Seed NL, Kaluarachchi ED (1998) Digital pulse interval modulation for optical communications. IEEE Commun Magaz 36(12):95–99
10. Ghassemlooy Z, Popoola WO (2010) Mobile and wireless communications network layer and circuit level design, chapter terrestrial free-space optical communications, pp 355–392. InTech
11. Nancy G, Dixit A, Jain VK, et al (2021) Capacity and BER analysis of BCH and LDPC coded FSO communication system for different channel conditions. Opt Quant Electron 53(5):1–25
12. Anshul J, Abaza M, Bhatnagar MR, Mesleh R (2020) Multipoint-to-multipoint cooperative multiuser SIM free-space optical communication: a signal-space diversity approach. IEEE Access 8:159244–159259
13. David G (1997) Forney Jr. on iterative decoding and the two-way algorithm. In: Proceedings of international symposium on turbo codes and related topics
14. Kiasaleh Kamran (2005) Performance of APD-based, PPM free-space optical communication systems in atmospheric turbulence. IEEE Trans Commun 53(9):1455–1461
15. Lange R, Smutny B, Wandernoth B, Czichy R, Giggenbach D (2006) 142 km, 5.625 Gbps free-space optical link based on homodyne BPSK modulation. In: Proceedings of SPIE 6105
16. Lee C-H, Wolf W (2005) Implementation-efficient reliability ratio based weighted bit-flipping decoding for ldpc codes. Electron Lett 41(13):755–757
17. Luong Duy A, Thang Truong C, Pham Anh T (2013) Effect of avalanche photodiode and thermal noises on the performance of binary phase-shift keying subcarrier-intensity modulation/free-space optical systems over turbulence channels. IET Commun 7(8):738–744
18. Mohan N, Ghassemlooy Z, Li E, Mansour Abadi M, Zvanovec S, Hudson R, Htay Z (2022) The BER performance of a FSO system with polar codes under weak turbulence. IET Optoelectron 16(2):72–80
19. Moradi Hassan, Refai Hazem H, LoPresti Peter G (2011) A switched diversity approach for multi-receiving optical wireless systems. Appl Opt 50(29):5606–5614
20. Richardson TJ, Urbanke RL (2001) Efficient encoding of low-density parity-check codes. IEEE Trans Inform Theor 47(2):638–656
21. Sangeetha RG, Hemanth C, Jaiswal I (2022) Performance of different modulation scheme in free space pptical transmission–a review. Optik 254:168675
22. Gaurav Soni and Jagjit Singh Malhotra (2012) Impact of beam divergence on the performance of free space optical system. Int J Sci Res Publ 2(2):1–5
23. Sugiyama H, Nosu K (1989) MPPM: a method for improving the band utilization efficiency in optical PPM. IEEE/OSA J Lightwave Technol 7(3):465–471
24. Wilson SG, Brandt-Pearce M, Cao Q, Leveque JH (2005) Frees-pace optical MIMO transmission with Q-ary PPM. IEEE Trans Commun 53(8):1402–1412
25. Youssef AA, Abaza M, Alatawi AS (2021) LDPC decoding techniques for free-space optical communications. IEEE Access 9:133510–133519
26. Yu K, Shu L, Fossorier MPC (2001) Low-density parity-check codes based on finite geometries: a rediscovery and new results. IEEE Trans Inform Theor 47.7
27. Zeidan HR, Elsabrouty MM (2008) Low complexity iterative decoding algorithm for low-density parity-check (LDPC) codes. In: 2008 1st IFIP wireless days, pp 1–5. IEEE

Optimization Techniques of DFIG Controller Design for Performance Intensification of Wind Power Conversion Systems

Om Prakash Bharti, Aanchal Verma, and R. K. Saket

Abstract This paper illustrates the optimization techniques of DFIG controller design for performance intensification of the wind power conversion system. The DFIGs are employed in wind energy conversion systems (WECSs) due to robustness toward changeable wind and rotor speed. DFIG kept the adaptable property since the system parameters are allocated with, including real, reactive power, DC-link voltage, and transient and dynamic responses. The analysis becomes more prominent during any unusual condition in the electrical power generation system. Therefore, the improvement in the system parameters for steady state and transient response performance of DFIG are required that can be accomplished using some controlling techniques. For fulfilling the task, the present work implements and compares the optimization methods for the design of the DFIG controller for WECS. The bio-inspired optimization techniques are applied to get the optimal controller design parameters for DFIG-based WECS. The optimized DFIG controllers are then used to repossess the transient response performance of the six-order DFIG model with a step input. The results using MATLAB/Simulink show that the firefly algorithm (FFA) over other control techniques has best performance as compared with the other controller design methods.

Keywords Doubly-fed induction generator (DFIG) · Induction generator (IG) · Wind turbine (WT) · Transfer function (TF) · Wind energy conversion systems (WECSs) · Proportional · Integral and derivatives (PID)

O. P. Bharti
Department of Electrical Engineering, Government Polytechnic College, Ghazipur, UP, India

A. Verma · R. K. Saket (✉)
Department of Electrical Engineering, Indian Institute of Technology (BHU), Varanasi, UP, India
e-mail: rksaket.eee@iitbhu.ac.in

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_35

1 Introduction

The demand of the electrical power is necessary for accomplishing to the advancement of all the countries. Renewable energy sources take over the traditional sources of generation of electrical power. The prime causes to utilize renewable power are the depletion of earth's energy resources and greenhouse gas emissions (GHG) [1–3]. Wind power generation is most important favorable objective of the modern renewable energy systems. It is pure and sufficient to produce no GHGs, which makes it is a fast-growing resource [4, 5]. The global universal aggregate capacity of wind power was 496 GW in 2017, whereas in 2018 it is 597 GW [6]. Extraction of wind energy is sperformed done through a WECS; Fig. 1 describes a simplified diagram of WECS based on the DFIG. This DFIG system is associated with a dynamic voltage restorer (DVR) for fragility of the controller, which is presented in Fig. 1 [7, 8]. The FFA-based controller has the best performance in the comparative study of all the designed controllers. For the robustness against the grid faults [9], DVR compensates for the fault line voltage and fulfills grid requirement to the DFIG system for normal working from any distributed load in the grid. According to the speed control norm, WECS are two types, i.e., permanent and variable speed. For efficiency seizing and maximum power, a variable velocity WECS is implemented. The wind speed is volatile, and to deal with it, a DFIG is working that constitutes a wound rotor induction generator (WRIG). The advantages of DFIGs are pointed out in Table 1. DFIG performs in sub, super, and synchronous modes corresponding rotor velocity. The working approaches of DFIG are accountable to real power (P), reactive power (Q) direction, at grid frequency by means of invariable DC-link voltage [10]. The DC-link voltage, real power, reactive power, transient operation, and dynamical performance are described in [11–14]. So it becomes obligatory to extend the control schemes intended for WECS based on DFIG. Simultaneously system parameters' dynamic as well as transient performances are in limits, which is helpful to investigate.

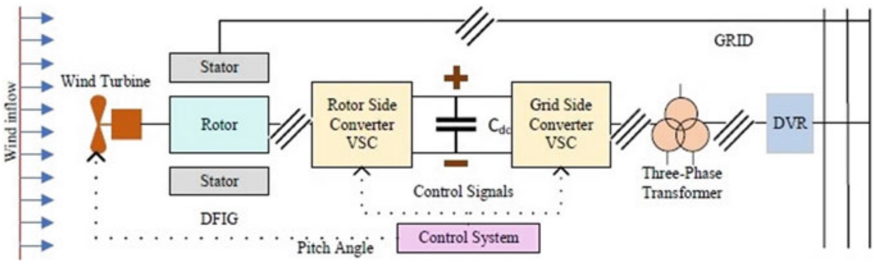


Fig. 1 Schematic diagram of WECS

Table 1 FFA-based controller gain

Controller gains	K_P	K_I	K_D
FFA-based	14.9800	139.9345	0.0009

The super-twisting conventional sliding mode control (SMC) and PI controllers are used to improve the dynamic performances and aerodynamic torque for WECS [15–17]. The $H\infty$ and SMC controller are utilized for the duration of the voltage dip [18]. The functions of H∞ have been observed in [19, 20], which describe how the proposed controller solves the mathematical optimization problems. To the fast grid synchronization and maximum power point tracking (MPPT), multi-objective based model predictive control (MPC) is applied in the WECS-based DFIG. The papers discussed previously could not be able to develop a comparative analysis of DFIG-based WECS controller design using bio-inspired or other control techniques. In this regard, this work describes optimization techniques to obtain optimal controller design for the DFIG-based WECSs.

The prime contributions of this manuscript are follows.

(i) A short learning on DFIG-based WECS to analyze the reactive power, active power, DC-link voltage, and step response of six-order DFIG model.
(ii) With accomplishment of static output feedback (SOF), bacterial foraging optimization (BFO), particle swarm optimization (PSO), firefly algorithm (FFA), genetic algorithm (GA), and differential evolutionary algorithm (DE) to achieve the optimization of the parameters of the WECS controller.
(iii) Assessment of the output responses to get enhanced method. The experimental results of FFA are better in this manuscript.

The comprehensive organization of the remaining manuscript is as follows. A brief study on WECS is illustrated in Sect. 2. DFIG controller design methodologies such as PSO, FFA, GA, BFO, and DE are discussed briefly in Sect. 3. Results with discussions are well given in Sect. 4. Lastly, the manuscript is finished with conclusion aspects in Sect. 5.

2 Wind Energy Conversion System (WECS)

As the name suggests, WECS is an energy conversion system. By this wind energy is converted into mechanical energy, and then, mechanical energy is converted into electrical power. The system arrangement is illustrated in Fig. 1. The governing mechanical power output of a WECS is described in expression (1). WECS consists of WT, gearbox, generator, voltage source converters, and a power transformer. A DFIG as a generator is the primary concern of this paper, and hence, it is elaborated on in subsequent subsections.

$$P_m = \frac{1}{2}\rho A_r v^3 C_p(\lambda, \beta) \tag{1}$$

P_m Mechanical energy collected from the rotor of WT,
ρ Density of air (kg per m^3),
A_r Swept area enclosed by WT blade (m^2) $= \pi . R^2$,

R Radius of WT blades (m),
C_p Performance improvement coefficient,
β Pitch angle,

$$\text{The tip's ratio of speed: } \lambda = \frac{R\omega_r}{v}$$

ω_r Rotational speed of wind turbine,
V velocity of wind.

3 DFIG Controller Design Techniques

Electric grid control is a significant part of an electrical power network operator, and it gets supplementary significant when refers to DFIG-based WT [20]. Wind speed deviation is noticeable, and the uncertain consumers' load is required to be controlled appropriately. A number of researches have been accomplished in building rugged and fully controllable DFIG-based WT integrated electrical grids. For optimum controller design, some optimization methods are utilized [19, 20]. Then six-order DFIG transfer function model is used to study the transient behavior parameters as given in Eq. (2). A six-order transfer function is used in this manuscript. As the mathematical modeling of the DFIG has, electric grid control is a significant part of an electrical power network operator. This gets additional momentous when it refers to DFIG-based power generation. The mathematical modeling of the DFIG has 6×6 state-space model matrix. The supervisory PID controller is well compared with various methods to ensure steady-state performances, such as zero steady-state error and state-space model matrix. The supervisory PID controller is well compared with various methods to ensure steady-state performances, such as zero steady-state error [18–20].

$$\begin{aligned}\text{T.F.} = {} & [0.000324S^6 - 1.75S^5 - 2366S^4 + 7.9 \times 10^6 S^3 \\ & + 7.5 \times 10^9 S^2 + 5 \times 10^{12} S + 2.18 \times 10^{14}]/ \\ & [S^6 + 2340S^5 + 8.67 \times 10^6 S^4 + 4.79 \times 10^9 S^3 \\ & + 2.7 \times 10^{12} S^2 + 1.27 \times 10^{14} S + 9.6 \times 10^{14}] \end{aligned} \tag{2}$$

3.1 Static Output Feedback (SOF)

For making the system stable, numerous controller design techniques are available. A SOF technique is the one that makes the system globally stable. It is required to obtain the adequate condition for the existence of SOF gains. For this, the focal contribution consists in providing adequate fresh conditions. The design of the SOF controller [20] is dependent on linear matrix inequality (LMI). LMI conditions are resulting for the SOF controller, which ensures the $H\infty$ performance and pole placement in the LMI contour. The designed SOF-based controller is implemented in DFIG system model for performance analysis.

3.2 Particle Swarm Optimization (PSO)

The technique is a new sequential computational scheme since 1995. By PSO scheme, the comprehensive solution is enhanced iteratively under some specified constraints. The PSO algorithmic flowchart is publicized in [20]. The PID controller gain is resulting by the fitness function of PSO, and this fitness function is implemented in the sixth-order DFIG model which results in the best fittest solution. PSO has optimization tool associated with swarm behavior, which depends on fish schooling or bird flocking behavior for supervising the particles to get the global best solutions. Its working methodology is accredited to the birds' flocking action. Positions of bacteria are shown, which move in clockwise and anticlockwise directions. The designed PSO-based controller is implemented in DFIG system model for performance analysis.

3.3 Bacterial Foraging Optimization (BFO)

The impression and brief description for algorithmic features-based bacterial foraging optimization (BFO) are illustrated in [19]. It combines bacteria and swarm's optimization such as bacteria chemotaxis, swarming, reproduction, in addition to elimination along with dispersal algorithm, respectively. It is motivated through the foraging activity of group bacteria, especially E. coli and M. Xanthus. The BFO is encouraged by the bacteria's chemotaxis response. However, the BFO algorithmic procedure as a flowchart and a graphical representation of swim as well as tumble for a bacterium is shown in [20]. The designed BFO-based controller is implemented in DFIG system model for performance analysis.

3.4 Genetic Algorithm (GA)

It is known that the optimization studies are divided into constrained and unconstrained problems. These problems are mainly based on the biological evolution process. The GA frequently updates the population of particular solutions. The individuals are selected randomly by GA at each step. The selected individuals are treated like parents, who are used to developing their children for upcoming steps or generations. By doing these successive generations, the population involves, and an optimal solution is obtained. The overall algorithm is described in [20]. The algorithm shows that the selection, combination, and mutation (SCM) are the main fitness assessment process. The designed GAO-based controller is implemented in DFIG system model for performance analysis.

3.5 Differential Evolutionary (DE)

DE is a meta-heuristic technique and uses the solution as the GA. DE is also a population-based algorithm as GA, which uses similar operators, as GA. This evolutionary method iteratively achieves an optimized solution. It has the property of convergence, obtaining global minima, and utilizing a few controlling parameters. The optimized solution is obtained by sustaining a population candidate's solution and creating a new candidate solution with active joins ones, as shown in [19, 20]. Thus, the problem is considered a black box that hardly provides the measured quality for a particular candidate solution. The designed DEO-based controller is implemented in DFIG system model for performance analysis.

3.6 Fire Fly Algorithm (FFA)

This algorithm is developed as a meta-heuristic, swarm intelligence, and naturally motivating technique. FFA works on a firefly's flashing behavior to attract partners, warning signal for predators, and establish communication among other flies. Its fitness function is dependent on the flashing intensity. The intensity is inversely proportional to the area of illumination. The mathematics shown below is utilized in determining an efficient, optimized solution. The flowchart of this algorithm is demonstrated in Fig. 2 which is implemented in this paper. FFA is a nature-inspired algorithm, which is based on the flashing light behavior of fireflies.

There are three rules for FFA, which are described below.

1. All fire flies are unisex and move toward brighter ones for their sex.
2. The firefly brightness is directly proportional to the degree of their of attraction, that decreases the distances from the other fire flies.
3. The attractive firefly moves randomly.

For optimization assessment, the fitness function is allied with the flashing light fitness function to get proficient optimal solutions. For searching solutions, the fireflies utilized two primary procedures that are: (i) Attractiveness (ii) Movement.

(i) **Attractiveness**: The attractiveness function of a firefly is associated with monotonically decreasing function as described in the following Eq. (3)

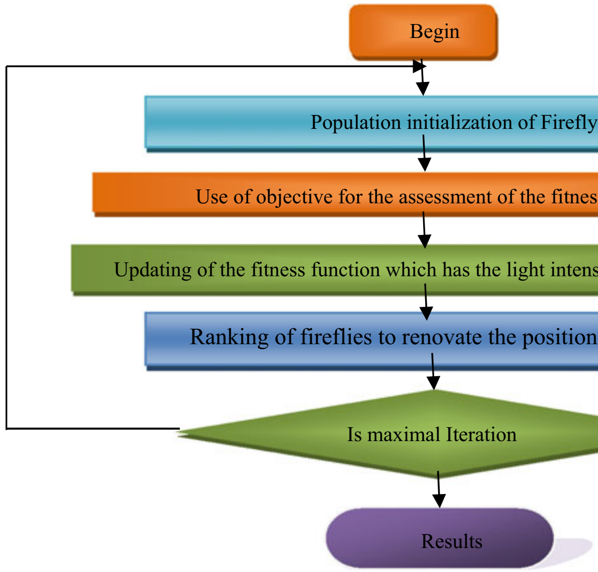


Fig. 2 Flowchart of the FA-based controller design

$$\beta(r) = \beta_0 e^{-\gamma r^m}; \; m \geq 1, \tag{3}$$

In this equation, r = distance measured between any two adjacent fireflies, β_0 = starting attractiveness of fireflies at $r = 0$, and γ = absorption parameter that controls the decrease of the light intensity. The distance r between any two ith and jth fireflies at position x_i and x_j, correspondingly, can be determined as an Euclidean or Cartesian as described in Eq. (4).

$$r_{ij} = \sqrt{\sum_{k=1}^{d} \left[x_{i,k} - x_{j,k}\right]^2} \tag{4}$$

where $x_{i,k}$ = kth spatial coordinate component x_i of the ith firefly, and d is the number of the dimension.

(ii) **Movement**: Equation (5) gives the movement description of a firefly; i is seduction by a brighter firefly j.

$$x_i = \left[x_i + \left\{\beta_0 \times e^{-\lambda r_{i,j}^2}\right\} \times (x_j - x_i) + \alpha\left(\text{rand} - \frac{1}{2}\right)\right] \tag{5}$$

Here, the primary term is the latest position of a firefly, and the second time is used for taking into consideration of the firefly's attractiveness toward the light intensity observed by adjoining fireflies. Moreover, the third term in the above equation is used for the arbitrary movement of a firefly. The parameter randomization is described by a coefficient and resolute by the difficulty of interest, whereas rand is a random numeral generator, which is uniformly distributed in the space [0,1]. Firefly algorithm for controller design is given below.

Step I: Initialize algorithm parameters,

Step II: For objective function $f(x)$ definition, where $x = (x_{1,} x_{2,} x_{3,} \dots x_d)^{\mathrm{T}}$,

Step III: For initial generation of population of fireflies or x_i ($i = 1, 2\dots, n$),

Step IV: For determination of the light intensity of I_i at x_i via f (x_i),

Step V: While ($t <$ maximum generation) for $I = 1, 2, 3, 4\dots$ up to n, (all n fireflies),

For $j = 1, 2, 3, 4\dots$ up to n, (n is number of fireflies) if ($I_j > I_i$), move firefly i toward j; end,

Step VI: If attractiveness varies with distance r via $\mathrm{e}^{-\lambda r^2}$,

Step VII: Compute new solution(s) and update intensity of light; End for j; End for i,

Step VIII: To give rank to fireflies and find the latest favorable result; End while,

Step IX: Post phenomenon/process results and visualization,

Step X: End the procedure.

The chosen fitness functions is described in equation (6)

$$F = (1 - e^{-6})(M_{\mathrm{p}} + E_{\mathrm{ss}}) + e^{-6}(T_{\mathrm{s}} - T_{\mathrm{r}}) \tag{6}$$

where M_{P} = Peak overshoot, F = Fitness function, E_{ss} = Steady-state error, T_{r} = Rise time, T_{s} = Settling time, and β = Scaling factor, which depends on the choice of the architect. For such case design, scaling factor $\beta = 1$. The MATLAB library has to define the PID parameters as a fitness function as values of input, and it returns the PID-based controlled model's fitness value as its output. The assumed fitness function is given in equation (7).

$$\text{Function}(F) = \text{fitness}[K_{\mathrm{D}}, K_{\mathrm{P}}, K_{\mathrm{I}}] \tag{7}$$

The PID parameters are the fitness function of the input and return the calculated fitness function value for different cases. The main objective is to minimize the fitness function as possible by associating other values of PID parameters.

Now, the controller gains for SOF, PSO, BFO, GA as well as DE-based have been shown in Table 2.

Table 2 Controller gains for various controller design techniques

Controller design techniques	K_P	K_I	K_D
PID reference	0.4635	6.7122	0.0009
SOF-based	0.0814	4.6647	0.0037
PSO-based	39.9781	7.6902	0.0271
BFO-based	0.1417	0.1472	0.1005
DE-based	4.7269	137.9634	0.0088
GA-based	0.7270	102.4343	0.0023

4 Results and Discussion

The detailed model of 9 MW WF with variable speed WT and associated DFIG is described in Sect. 4. The DFIG output waveforms and DFIG sixth-order T.F. system model transient response for the SOF, PSO, BFO, GA, and DE-based controller are shown in Figs. 3a–c and 4a, b. Transient performance parameters for various controller design techniques are shown in Table 3.

It is illustrated in Fig. 3a–c that the outputs that DFIG waveforms are enhanced positively by utilizing the FFA technique over all other optimization techniques. On the other hand, Fig. 4a, b and Table 3 depict that FFA-based design controller transient performance in terms of rise time, peak time, overshoot, settling time, undershoot, and peak value have been more optimized and improved.

Now, comparing various control techniques is implemented on MATLAB R2017b using Simulink. Tables 1, 2, and 3 show the comparative analysis. Tables depict controller gain values and performance parameters by the implemented controller

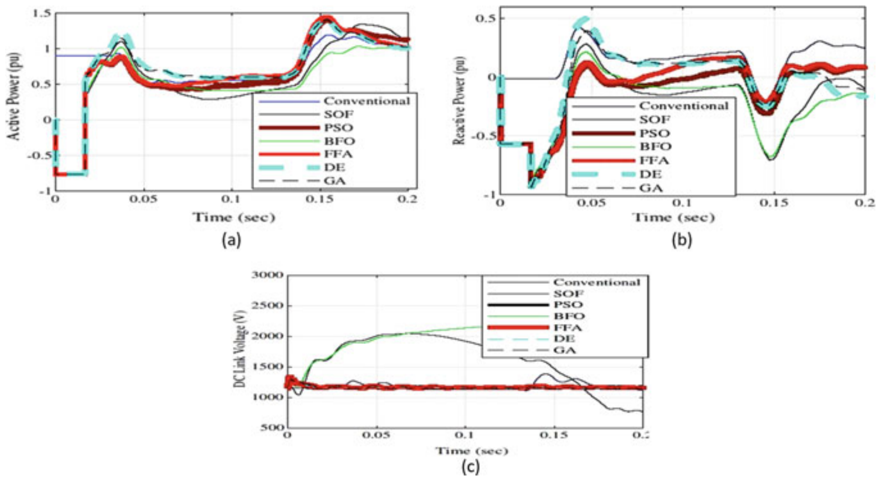


Fig. 3 **a** Active power comparison of all methods. **b** Reactive power comparison of all method. **c** DC-link voltage comparison of all methods

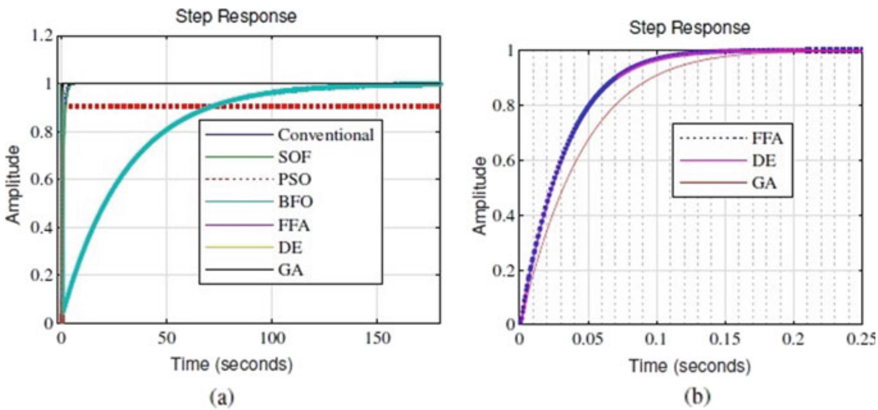


Fig. 4 **a** Step response comparison of all methods. **b** Step response of FFA, DE, and GA

Table 3 Transient performance parameters for various controller design techniques

Controller design techniques	T_R(s)	T_S(s)	T_P(s)	Overshoot	Undershoot	Peak value
PID reference	0.0153	0.5292	0.1420	15.1986	8.1861	1.1513
SOF-based	0.2134	0.5277	1.0000	0.0000	1.5475	0.9997
PSO-based	0.094	0.4062	0.0000	0.0000	13.8755	0.5000
BFO-based	35.02	62.3542	0.0000	0.0000	0.0000	0.5000
FFA-based	0.0660	0.1107	0.2104	0.0001	0.0000	1.0000
DE-based	0.0676	0.1146	0.2116	0.0000	0.0000	0.9993
GA-based	0.0912	0.1541	0.2541	0.0000	0.0000	0.9991

technique. The following comparisons are made in the subsequent subsections of this paper to show the enhancement of the proposed control technique.

(i) The active power of the conventional controller is compared with SOF, PSO, BFO, FFA, DE, and GA-based controllers, as shown in Fig. 3a.
(ii) Then, the reactive power output of a conventional controller is compared with SOF, PSO, BFO, FFA, DE, and GA-based controllers, as mentioned in Fig. 3b.
(iii) DC-link voltage output is analyzed with all control techniques in Fig. 3c.
(iv) Figure 4a, b analyzes the dynamic performance of various methods to obtain the best controller design technique.

The step responses of these three controller design techniques are shown in Fig. 4a, b. The responses are then compared with a conventional controller. A comparison between DE, GA, and FFA step responses by obtaining the transient performance parameters is described in Table 3. It is concluded from the table that peak, settling, and rise times with GA > DE > FFA, respectively. It can be observed from the responses that the implemented techniques DE and GA enhance the system responses.

On the other hand, FFA not only improves the output of the system response but also reduces the overshoot to zero. As a result, it is depicted that the FFA technique provides a much better way to achieve a reliable and efficient controller for DFIG-based WT.

5 Conclusion

In this research work, FFA algorithmic approach is implemented for controller design. This controller is utilized in DFIG-based WT system model for performance analysis. The SOF-, PSO-, BFO-, GA-, and DE-based designed controllers are studied for DFIG-based WECS system operation. Also, the transient performance in terms of rising time, peak overshoot, and settling time have been improved by using soft computational evolutionary optimization method. Further, the techniques are irrespective of the type of parameter sensitivity, evade local optimization solutions. Finally, it is observed that the FFA technique is suitable in comparison with other control techniques. It can be seen that the peak time, settling time, and rise time for FFA are found to be lesser than the DE- and GA-based control techniques. Also, FFA-based designed controller reduces overshoot to zero and enhances the system response.

Acknowledgements The researchers are heartily thankful to electric machines and drive complex, and control systems complex of EED of IIT (BHU) India for the laboratory facilities to complete this manuscript for Springer Nature conference. This research work is vigorously dedicated to the first author's baby doll 'Ananya Singh'.

References

1. Bihari SP, Sadhu PK, Sarita K, Khan B, Arya LD, Saket RK, Kothari DP (2021) A comprehensive review of micro grid control mechanism and impact assessment for hybrid renewable energy integration. IEEE Access, 9:88942–88958
2. Council, Global Wind Energy (2021) GWEC global wind report 2021. Global Wind Energy Council, Brussels, Belgium
3. Musarrat MN, Fekih A, Islam MR (2021) An improved fault ride through scheme and control strategy for DFIG-based wind energy systems. IEEE Trans Appl Supercond 31(8):1–6
4. Angala Parameswari G, Habeeb ullah Sait H (2020) A comprehensive review of fault ride-through capability of wind turbines with grid-connected Doubly Fed Induction Generator. Int Trans Electric Energy Syst 30(8):e12395
5. Kong X, Wang X, Abdelbaky MA, Liu X, Lee KY (2022) Nonlinear MPC for DFIG-based wind power generation under unbalanced grid conditions. Int J Electr Power Energy Syst 134:107416
6. Peng X, Liu Z, Jiang D (2021) A review of multiphase energy conversion in wind power generation. Renew Sustain Energy Rev 147:111172
7. Bharti OP, Saket RK, Nagar SK (2019) Reliability assessment and performance analysis of DFIG-based WT for wind energy conversion system. Int J Reliab Saf 13(4):235–266

8. Huang S, Wu Q, Guo Y, Rong F (2019) Hierarchical active power control of DFIG wind farm with distributed energy storage based on a DMM. IEEE Trans Sustain Energy
9. Verma B, Padhy PK (2018) Optimal PID controller design with adjustable maximum sensitivity. IET Control Theor Appl 12(8):1156–1165
10. Jigang H, Hui F, Jie W (2019) A pi controller optimized with modified differential evolution algorithm for speed control of BLDC motor. Automatika 60(2):135–148
11. Tilli A, Conficoni C (2019) An effective control solution for doubly-fed induction generator under harsh balanced and unbalanced voltage sags. Control Eng Pract 84:172–182
12. Hu J, Li Y, Zhu JG (2018) Multi-objective model predictive control of doubly-fed induction generators for wind energy conversion. IET Gener Trans Distrib 13(1):21–29
13. Bharti OP, Sarita K, Aanchal Singh S, Vardhan AS, Vardhan S, Saket RK (2021) Controller design for DFIG-based WT using gravitational search algorithm for wind power generation. IET Renew Power Gener 15(9):1956–1967
14. Moghadam FK, Ebrahimi SM, Oraee A, Velni JM (2018) Vector control optimization of DFIGs under unbalanced conditions. Int Trans Electr Energy Syst 28(8):2583
15. Jiang T, Zhang Y (2021) Robust predictive rotor current control of doubly fed induction generator under unbalanced and distorted grid. IEEE Trans Energy Convers
16. Pura P, Iwański G (2021) Rotor current feedback based direct power control of a doubly fed induction generator operating with unbalanced grid. Energies 14(11):3289
17. Bharti OP, Saket RK, Nagar SK (2017) Controller design for doubly fed induction generator using particle swarm optimization technique. Renew Energy 114:1394–1406
18. Bharti OP, Saket RK, Nagar SK (2016) Controller design for DFIG driven by variable speed wind turbine using static output feedback technique. Eng Technol Appl Sci Res 6(4):1056–1061
19. Bharti OP, Saket RK, Nagar SK (2017) Controller design of DFIG based wind turbine by using evolutionary soft computational techniques. Eng Technol Appl Sci Res 7(3):1732–1736
20. Bharti OP, Saket RK, Nagar SK (2018) Controller design of DFIG-based WT by using de-optimization techniques. In: 2018 SICE international symposium on control systems (SICEISCS). IEEE, pp 128–135

The Relationship Between Social Media Influencers (SMIs) and Consumers' Purchase Behaviour in Malaysia

Tang Mui Joo and Chan Eang Teng

Abstract Social media influencer marketing has become a crucial marketing strategy of a company upon the development of social media with the increasing number of social media users. This research is to study whether the characteristics of an influencer are important in affecting the consumers' purchase behaviour. This research is also to identify the relationship between SMIs and influencer marketing in affecting consumers' purchase decision. Lastly, it is to determine the media strategies that SMIs used to influence consumers' purchase behaviour. Two-step flow theory and electronic word of mouth have been discussed as to look at how social media influencers' marketing strategies are affecting consumers' purchase decision. Online survey is used. The subjects for this research are 150 volunteering adults, aged between 18 and 24 years old with the criteria that the subjects must have at least been following three influencers on social media and have noticed the influencer endorsement on social media before. Google Form has been used, and snowballing is used to reach the subjects. This research has concluded that the positive characteristics of SMIs are bringing effects in consumers' purchase decision. It is also found that social media is the effective platform for most of the respondents to seek product information. There comes along with the rise of SMIs who play the role of opinion leaders strategizing with EWOM that manage to influence consumers' purchase decision. Though so, friends and family are still playing an important role in the purchase decision.

Keywords Social media influencers · Characteristics of SMIs · Consumers' purchase decision · Consumers' purchase behaviour · Two-step flow theory · Electronic word of mouth

---

T. M. Joo (✉) · C. E. Teng
Tunku Abdul Rahman University of Management and Technology, 53300 Kuala Lumpur, Malaysia
e-mail: muijoo@hotmail.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_36

1 Introduction

Social media influencer marketing has become a crucial marketing strategy of a company upon the development of social media with the increasing number of social media users. Social media influencers (SMIs) are an online advocate to influence their audience by expressing their experiences towards the certain products or services which they encounter in their daily lives. The development of social media enables individuals who have real-time experience of a particular product or service to share their opinions and discuss with others, thus becoming a vital and powerful input for people to make purchasing decisions. As the need for a consumer's view and its ability to affect other consumers has increased, SMIs have accordingly arisen [1].

In Malaysia as of January 2022, there were 30.25 million social media users which was 91.7% of the total population in the country [2]. Looking at the rise of social media users and SMIs, the consumer behaviour will also be changed by SMIs [3].

The purposes of this research have been as below:

(a) To study whether the characteristics of an influencer is important in affecting the consumers' purchase behaviour
(b) To identify the relationship between SMIs and influencer marketing in affecting consumers' purchase decision
(c) To determine the media strategies that SMIs used to influence consumer purchase behaviour.

To achieve the purposes of this research, this paper discusses the characteristics of influencers that affect consumer purchase behaviour. The characteristics include credibility [4], trustworthiness, and expertise in influencers that will affect the consumer purchasing behaviour [5]. This paper then discusses on the relationship between influencer marketing (IM) and social media influencers (SMIs). The relationship shall indicate the effect of SMIs on consumers' purchase behaviour. Further to it, this paper will also determine the marketing strategy used by SMIs based on two-step theory and electronic word of mouth (EWOM) in the approach to the consumers' purchase behaviour.

For data collection, voluntary online survey and snowballing have been used upon the samples of this research who are adults aged between 18 and 24. The criteria set in the survey question are that the respondents shall have followed at least three influencers on social media and have noticed the influencer endorsement on social media before. Conclusion is then drawn from the data collected. It is concluded that the positive characteristics of SMIs are bringing effects in consumers' purchase decision. It is also found that social media is the effective platform for most of the respondents to seek product information. There comes along with the rise of SMIs who play the role of opinion leaders strategizing with EWOM that manage to influence consumers' purchase decision.

2 Literature Review

2.1 *The Characteristics of SMIs in Influencing Consumers' Purchase Behaviour*

The advent of technology and social media has promoted the rise and then increase of SMIs, and more marketers are utilizing SMIs to boost their sales [6]. SMIs have established themselves as endorsers by generating a range of buzzwords, and they are deemed to be the cost-efficient and effective marketing trends compared to other marketing strategies [7]. SMIs can change the purchasing behaviour of consumers who follow them and make them take their advice, when the SMIs are trusted [8].

There are several characteristics of SMIs which include credibility, trustworthiness, and expertise. Credibility is a positive characteristic of the endorser that affects the message he transmits to the receiver [9]. It can be defined as whether a person is identified genuine, impartial, and factual [10]. The SMIs are credible and influential because of their lifestyle reflected in vlog or blog, and they update their daily activities to their followers. SMIs communicate with their followers in a real and true way which leads them to interact and participate in the content of SMIs by posting, liking, commenting, and sharing the content to their people surrounding them [11]. However, this will only happen if the audience has a positive perception of the SMIs [8]. As a consumer, they believe that the credible information is the main factor in the purchase decision. However, if the consumers feel that the information is partial and not authentic, then the credibility of the SMIs will be decreased [10]. Credibility is expected to lead to more positive attitudes towards both the endorser and the endorsement [12]. A highly credible SMIs would cause the consumers to believe that the recommended product or service was actually to boost their own image, communicate their genuine interest in the product, or convey their intention to help others [8].

Credibility model has been measured using two subcomponents which are trustworthiness and expertise [12]. Trustworthiness means that the extent of confidence of the consumers in receiving the information from influencers. The number of followers is an important factor to determine the trustworthiness of an influencer; the higher level of trustworthiness of an influencer will be more persuasive to the followers [9]. Consumers tend to trust online messages shared by opinion leaders [13]. It is also found that followers trust the branded posts delivered by influencers may positively affect their purchase behaviour [14]. As for the expertise, the influencers are perceived to possess the relevant knowledge, skills, and practices to promote the product [4]. When the influencers shape themselves as an expert within a certain market, their influencers brand personality will match the certain products endorsed within their market naturally [12] such as "food vlogger", "fitness vlogger", "sport blogger" and "make-up vlogger" and keep updating the product or brand information [15].

It is concluded here that the credibility, trustworthiness and expertise are closely related and are also the vital components to the success of a SMI. When the consumer makes a purchase decision, they will observe the characteristics of SMIs. The characteristics of SMI are playing a role in influencing the consumers' purchase behaviour and decision.

2.2 The Relationship Between Social Media Influencers (SMIs) and Influencer Marketing (IM)

Social media influencers (SMIs) are an online celebrity with a large portfolio of followers on one or more social media platforms such as YouTube, Facebook, Instagram and Blogs in affecting their followers. SMIs are ordinary people who have become online influencers via social media platforms creating and sharing contents, as opposed to traditional celebrities who are famous through traditional media such as film and television shows [14]. SMIs have a great attractiveness towards brands because they are viewed by consumers as personal, authenticity, credibility and solid source of information with the remarkable benefit of a large audience for the brands [16].

Influencer marketing (IM) is a marketing strategy with the influence of opinion leaders or key individuals to drive consumers' brand awareness and their purchase decisions [14]. Influencers can be anyone as long as they are able to influence others in a particular industry and community as well as to encourage people to try and use their products or services based on their suggestions [16].

Social media platforms are the dominant tool for people to exchange information and build relationships [14]. Therefore, IM has become an up-to-date and efficient brand marketing strategy that marketing managers are interested in [17]. It focuses on utilizing SMIs to drive brand awareness via social media in order to reach a huge target audience, and it also became an important promotional tool with a significant impact on consumer purchase behaviour [18].

The relationship is reflected in this way that IM is closely related to social media platforms where the influencers are social media advocates who are called as SMIs. SMIs play an important role in influencer marketing in influencing the consumers' purchase behaviour and decision.

2.3 Two Steps Flow and Electronic Word of Mouth (EWOM) in SMIs

Two-step flow theory predicts that media indirectly influences the public via information opinion leaders who deliver messages on the network [19]. The two-step flow theory shows that information from the media may not always reach the general public right away. Instead, some people who are considered opinion leaders interpret and decipher the information they get from the media before presenting it to the audience. SMIs can be described as opinion leaders because they are the promoter of a brand or company [20]. They are influential and credible people from whom other people ask for suggestions, and they often have discussions with their "audience", so-called opinion followers [21].

A large amount of product purchase can only happen with the reviewing of others. At the same time, when the person wants to seek advice before purchasing, they may look to more than one person who are opinion leaders. For instance, YouTubers were evidently acting as opinion leaders, also in an influencer marketing context, and they had an effect on the views of the teenage girl on the fashion product [22]. There is a correlation between opinion leaders and consumers' purchase behaviour, and thus, the information delivered by SMIs on social media platforms has a great influence for consumers who seek for advice before purchasing [21]. SMIs can be described as opinion leaders in the two-step flow theory, and they are influential enough to influence their opinion followers. The impact will later affect the consumers' purchase behaviour effectively through social media.

On the other hand, looking at electronic word of mouth as a marketing strategy, it is extremely effective in forming and shaping consumer attitudes, behavioural, and purchase intentions [23]. People share or collect information from known individuals about the particular product or service before purchase decision [24]. EWOM can be any statement or recommendations about the negative or positive of a certain brand, product, or company that is spread by potential, actual or former customers through the Internet and also can be explained as a communication directed to consumers on the Internet [4].

The sharing of SMIs can also be considered EWOM as they represent users or buyers sharing their experience and evaluation of a certain product or service with other potential customers [25]. It is also stated that influencer marketing could be seen as an extension of WOM [8].

According to the latest data in January 2022, there were 29.55 million Malaysian Internet users which is the penetration rate of 89.6% of the total population in Malaysia [2]. With such high penetration of the total population, EWOM has had a great impact on people's lives, thereby leading to changes in consumers' purchase behaviour [18]. Consumers' purchase behaviour will be affected by EWOM of SMIs, since the feedback and sharing of SMIs are the reference of certain products for the consumers.

The relationship of two steps flow, EWOM, and SMIs can be simplified as such that with the characteristics of a SMI, he/she plays the role as an opinion leader. He/she may be utilizing EWOM as a marketing strategy in influencing the consumers' purchase decision.

3 Methods

The researchers use quantitative design and online survey to collect data. Online survey has been chosen as it allows the researchers to collect data from a big number of respondents at no geographical barrier. Online survey is at a zero cost. There is no face-to-face contact needed. The replies from respondents can be much more instant. The questionnaire in the format of Google Form will be distributed through emails and social media platforms to targeted respondents though snowballing among volunteered friends and families. The questionnaire is divided into three sections. Section one is on demography, section two is on the effects in characteristics of SMIs on consumers' purchase decision, and section three is on the relationship between SMIs, influencer marketing and consumers' purchase decision. All the questions are close-ended and partly Likert. The time frame of the survey is a month, starting from 1 August 2022 to 31 August 2022.

The subjects are 150 adults, aged within 18 to 24 years old based on the rational that this age range is the highest social media users [26]. The subjects must have at least been following three influencers on social media and have noticed the influencer endorsement on social media before. This is a voluntary-based online survey using Google Form and snowballing from the volunteers.

4 Results

4.1 Demographics

Table 1 displays a detailed summary of the demographic profiles of the 150 respondents. The demographic profile includes the respondents' gender, age group, ethnicity, and income level.

The majority of the respondents are between the ages of 19–21, which are 56% over 150 respondents. There are 35 respondents belonging to the age group between 22 and 24 years old and 31 respondents from the age group of 15–18. There are 61.3% of female respondents over 38.7% of male respondents. Among the ethnicity, 64% of the respondents are Chinese, 21.3% of Malays and 14.7% of Indians.

For the respondents' income level, 61 respondents' income level is between RM1 and RM5000, and 44 of the respondents have no income. There are 18% of the respondents at the income level of RM501–RM1000, whereas respondents who have income between RM1500 and above are 8% of the respondents. There are only 4% of the respondents from the income level RM1000–RM1500.

Table 1 Demographic profile

Demographic profile	Number	Percentages (%)
Age		
15–18	31	20.7
19–21	84	56
22–24	35	23.3
Gender		
Male	58	38.7
Female	92	61.3
Ethnicity		
Chinese	96	64
Malay	32	21.3
Indian	22	14.7
Income		
No income	44	29.3
RM1–RM500	61	40.7
RM501–RM1000	27	18
RM1000–RM1500	6	4
RM1500–above	12	8

Source Online survey conducted from 1 August 2022 to 31 August 2022

4.2 *The Effects in Characteristics of SMIs on Consumers' Purchase Behaviour*

This section entails the characteristics of SMIs that eventually bring effects on consumers' purchase behaviour (Table 2).

Among 150 respondents, the data shows that 66.7% of the respondents sometimes follow SMIs in social media platforms. Follow on with 22% of often, 9.3% of not

Table 2 Frequency of respondents in following SMIs on social media platform

How frequent do you follow social media influencers (SMIs) on social media platform?	Frequency	Percentage (%)
1. Never	3	2
2. Not often	14	9.3
3. Sometimes	100	66.7
4. Often	33	22

Source Online survey conducted from 1 August 2022 to 31 August 2022

often and 2% that never follow any SMIs. This question reflects on the number of opinion followers who follow SMIs for their latest news and update.

For the response on the question about how frequent the respondents purchase products because of the influencers, there are 55.3% of the respondents sometimes purchase because of the SMIs. There are only 20.7% of the respondents who have chosen “Often” and “Not Often” are 18.7%. Followed by only 5.3% of the respondents have never purchased. The responses reflect that SMIs have the strong characteristics in influencing the followers.

Table 3 reflects on the information sought by the respondents before making purchases.

Table 3 shows that before the respondents make any purchases, out of 150 respondents, 149 of them seek for products reviews and product ratings, following with 148 of them seek recommendation from family and friends, 144 of them seek for top selling product and lastly 140 of them seek for product repost. It is reflected that product review is important because product reviews help them get a clear idea of the product before making a purchase.

Follow on with the question about how much the respondents trust SMIs. There are 64.7% of the respondents trust SMIs, 6% trust less and 29.3% of them are neutral. People trust SMIs based on the credibility which is a positive characteristic of the influencers.

For this section, it is shown that the characteristics of SMIs in terms of their credibility, trustworthiness and expertise in product knowledge would attract opinions followers to decide on their purchase. That reflects on two-step flow theory that information from the media may not always reach the general public right away. Instead, some people who are considered opinion leaders interpret and decipher the information they get from the media before presenting it to the audience.

EWOM also plays a key role in people’s purchase decisions when buying a product. Consumers rely on information provided by others online to find authentic sources before making a purchase decision, exposing the quality and risks of products, which can profoundly affect their behaviours, attitudes, purchase intentions, and then purchase decisions.

Table 3 Information sought by the respondents before making purchases

As a consumer, what is the information you sought before making purchases?	Strongly agree (%)	Agree (%)	Disagree (%)	Strongly disagree (%)
Product reviews	88	61	0	1
Product ratings	71	78	0	1
Product repost	49	91	9	0
Top selling product	76	68	5	1
Recommendations from family and friends	64	84	1	1

Source Online survey conducted from 1 August 2022 to 31 August 2022

4.3 The Relationship Between SMIs and Consumers' Purchase Decision

This section would entail the relationship between SMIs and influencer marketing that eventually affect consumers' purchase decision.

The section starts with the question if the consumers like to be notified with the latest information regarding SMIs. 76% of them agreed, 2% of them disagree, and 22% are neutral on it. It shows that social media is an important tool for SMIs since the notifications from social media platforms are what SMIs cannot do without them. It is to show that the media strategies of two-step flow theory and EWOM need social media to support in influencing their followers.

The next question is about the types endorsements done by SMIs that may affect consumers' purchase decision. The responses ar4e summarized in Table 4.

The collected data shows that 83 respondents which are 55.3% of them choose "social media influencers conduct a Facebook/Instagram/YouTube/TikTok live to test and give reviews towards the certain product". Then, "social media influencers record a short reel/video to share their experience on the particular product" has the highest score of 60.7%. Next, 48.7% of the respondents have chosen "social media influencers indicate in the description that they are experts in a certain field (travel vlogger, fitness vlogger, make-up vlogger, etc.)" followed by 31.3% of the respondents who have chosen "social media influencers conduct a giveaway section (particular endorsement product: cosmetic, air tickets, fitness product)". Last but not least, 22% of the respondents have chosen "social media influencers create a hashtag of a certain product to have a linkage with product and audience".

Here it shows that the types of endorsement through different platforms may attract opinion followers on top of their credibility, trustworthiness and expertise

Table 4 Types of endorsements by SMIs that affect consumers' purchase decision

What do you think Social Media Influencers (SMIs) do on social media will make you change your mind about him starting to purchase as his/her endorsement?	Frequency	Percentages (%)
Social media influencers conduct a Facebook/Instagram/YouTube/TikTok live to test and give reviews towards the certain product	83	55.3
Social media influencers record a short reel/video to share their experience on the particular product	90	60.7
Social media influencers indicate in the description that they are experts in a certain field (travel vlogger, fitness vlogger, make-up vlogger, etc.)	73	48.7
Social media influencers conduct a giveaway section (particular endorsement product: cosmetic, air tickets, fitness product)	47	31.3
Social media influencers create a hashtag of the certain product to have a linkage with product and audience	33	22

Source Online survey conducted from 1 August 2022 to 31 August 2022

which are closely related and are important to determine the success of a SMI in influencing consumers' purchase decision.

5 Conclusion

As a conclusion, it is reflected from the respondents that the positive characteristics of SMIs are bringing effects in consumers' purchase decision. Those positive characteristics are credibility, trustworthiness and expertise. That reflects on two-step flow theory that information from the media may not always reach the general public right away. Instead, some people who are considered opinion leaders interpret and decipher the information they get from the media before presenting it to the audience. That is where the consumers may be seeking information from SMIs who are playing the roles of opinion leaders. On the other hand, in Malaysia, the recommendations of friends and family are one of the information they seek for before making a purchase decision. This study can strongly prove that consumers will not overly rely on the information on social media but also the feedback from the people around them.

EWOM also plays a key role in people's purchase decisions when buying a product. Consumers rely on information provided by others online to find authentic sources before making a purchase decision, exposing the quality and risks of products, which can profoundly affect their behaviours, attitudes, purchase intentions and then purchase decisions.

It is found that social media is the effective platform for most of the respondents to seek product information. There comes along with the rise of SMIs who play the role of opinion leaders strategizing with EWOM that manage to influence consumers' purchase decision. This research further implies the trend of social media marketing and purchase behaviour in Malaysia. For the future development, there is a potential to integrate social media marketing with other marketing strategies to accommodate the future market needs.

The authors acknowledged the raw materials from Cho Jia Ying, Lim Tze Yean, Teo Rou Jie, Tan Wei Ying and Wong Chen Xuan.

References

1. Anwar H, Gayathri A. https://www.researchgate.net/publication/362174282_SOCIAL_MEDIA_INFLUENCERS%27_MARKETING. Last accessed 20 July 2022
2. Kemp S. https://datareportal.com/reports/digital-2022-malaysia. Last accessed 25 July 2022
3. Zhang X, Ding X, Ma L (2022) https://sci-hub.hkvisa.net, https://doi.org/10.1080/0144929X.2020.1800820. Last accessed 21 July 2022
4. Kwiatek P, Baltezarevic R, Papakonstantinidis S. https://www.researchgate.net/publication/354663248_THE_IMPACT_OF_CREDIBILITY_OF_INFLUENCERS_RECOMMENDATIONS_ON_SOCIAL_MEDIA_ON_CONSUMERS_BEHAVIOR_TOWAR DS_BRANDS. Last accessed 23 July 2022

5. Balaban DC, Mucundorfeanu M, Naderer B. https://www.degruyter.com/document. Last accessed 28 July 2022
6. Wei L, Singh J, Kularajasingam J. https://ejbm.sites.apiit.edu.my/files/2022/01/Paper-5-Impact-of-Social-Media-Influencers-On-Purchasing-Intention-Towards-Pet-Products.-A-Quantitative-Study-Among-Females-in-Malaysia.pdf. Last accessed 14 Aug 2022
7. Jean L, Rozaini A, Radzol M, Hwa C, Wong M. https://www.researchgate.net/publication/330635364_The_Impact_of_Social_Media_Influencers_on_Purchase_Intention_and_the_Mediation_Effect_of_Customer_Attitude. Last accessed 15 Aug 2022
8. Zainab AD, Zahra AM, Shilan R. https://www.diva-portal.org/smash/get/diva2:1437746/FULLTEXT01.pdf. Last accessed 10 July 2022
9. Abdullah J. http://eprints.utar.edu.my/4026/1/fyp_AV_2020_IAAZ_-_1800437.pdf, Last accessed 20 June 2022
10. Grafström J, Jakobsson L, Wiede P. https://www.diva-portal.org/smash/get/diva2:1214105/FULLTEXT01.pdf. Last accessed 20 July 2022
11. Singh K. https://www.researchgate.net/publication/354636053_Influencer_Marketing_from_a_Consumer_Perspective_How_Attitude_Trust_and_Word_of_Mouth_Affect_Buying_Behavior. Last accessed 16 Aug 2022
12. Janssen L, Schouten AP, Croes E. https://www.tandfonline.com/doi/full, https://doi.org/10.1080/02650487.2021.1994205. Last accessed 15 Aug 2022
13. Pop RA, Săplăcan Z, Dabija DC, Alt MA. https://www.tandfonline.com/doi/full/10.1080/13683500.2021.1895729?casa_token=rYokycUWTsEAAAAA%3AcfKmHx27esUS0-ucrG3HGq0-GGIAzZx2R6DNG76bln5MZ13eWKj1flTUvU3ivDrEP1OPjIq EJ4fiuw. Last accessed 15 July 2022
14. Lou C, Yuan S (2022) https://www.tandfonline.com/doi/full, https://doi.org/10.1080/15252019.2018.1533501. Last accessed 15 Aug 2022
15. Chekima B, Chekima FZ, Adis A. https://deliverypdf.ssrn.com/delivery.php?ID=244094087111098084081095097095005122002041002072040074081007125108030102090087016092110057106062021112049088104007017022096084119017001015088023095105064077097090028020036007092110002084007127068026065013100112021077090001089110029065113011069027071005&EXT=pdf&INDEX=TRUE. Last accessed 17 July 2022
16. Harrigan P, Daly T, Coussement K, Lee J, Soutar G, Evers U. https://sci-hub.hkvisa.net, https://doi.org/10.1016/j.ijinfomgt.2020.102246. Last accessed 10 July 2022
17. Nguyen C, Nguyen T, Luu V (2022) Relationship between influencer marketing and purchase intention: focusing on Vietnamese gen Z consumers. Independent J Manage Prod 13(2):810–828
18. Liu H, Shaalan A, Jayawardhena C. https://www.researchgate.net/publication/362570328_The_Impact_of_Electronic_Word-of-Mouth_eWO. Last accessed 13 July 2022
19. Tyagi M, Kumar MD, Kumar M, Kumar P. https://www.journalppw.com/index.php/jpsp/article/view/8629/5640. Last accessed 28 July 2022
20. Watkins B. https://books.google.com.my/books?id=y_EfEAAAQBAJ&pg=PA25&dq=two+step+flow+influencer+marketing+text+book&hl=en&sa=X&ved=2ahUKEwiuj7jUltj5AhWd7zgGHVgRDw8Q6AF6BAgHEAI#v=onepage&q=two%20step%20flow%20influencer%20marketing%20text%20book&f=false. Last accessed 14 July 2022
21. Norhio E, Virkkunen P. https://www.diva-portal.org/smash/get/diva2:1321153/FULLTEXT01.pdf. Last accessed 8 Aug 2022
22. Leikas N, Szkwarek K. https://www.diva-portal.org/smash/get/diva2:1482544/FULLTEXT01.pdf. Last accessed 7 Aug 2022
23. Wegmann OP. https://research-api.cbs.dk/ws/portalfiles/portal/59790349/485616_Master_Thesis_Influencer_Marketing_digital_aflevering.pdf. Last accessed 28 July 2022
24. Rani A, Nagesh SH. https://www.researchgate.net/publication/345603959_Electronic_Word_of_Mouth_eWOM_Strategies_to_Manage_Innovation_and_Digital_Business_Model. Last accessed 14 Aug 2022
25. Hussain S, Song X, Niu B. https://www.frontiersin.org/articles, https://doi.org/10.3389/fpsyg.2019.03055/full. Last accessed 13 Aug 2022

26. Dwidienawatia D, Tjahjana D, Abdinagoro SB, Gandasari D, Munawaroh. https://www.sciencedirect.com/science/article/pii/S2405844020323860. Last accessed 23 July 2022
27. Ismail N, Ahmad J, Noor S, Jayslyn S (2019) Malaysian youth, social media following, and natural disasters: what matters most to them? Media Watch 10(3):508–521

HUM: A Novel Algorithm Based in Blockchain for Security in SD-WAN Controller

Jorge O. Ordoñez-Ordoñez, Luis F. Guerrero-Vásquez, Paul A. Chasi-Pesántez, David P. Barros-Piedra, Edwin J. Coronel-González, and Brayan F. Peñafiel-Pinos

Abstract Currently, software-defined networks (SDN) are displacing traditional networks, and this carries a lot of advantages and disadvantages, one of which is security. For this reason, this article presents a security analysis on wide area SDN network controllers. Furthermore, we propose the use of the HUM algorithm, a blockchain-based algorithm, as a possible solution to increase the robustness of security in the flow of packets between edge devices. This algorithm works by a group of controller nodes that are aware of all changes made to the data flow. The simulation of a topology is presented, and finally, an application case is proposed for the use of the algorithm within an SD-WAN network in a financial institution.

Keywords SDN · SD-WAN · Security · Blockchain · Encryption · Cybersecurity

J. O. Ordoñez-Ordoñez (✉) · L. F. Guerrero-Vásquez · P. A. Chasi-Pesántez · D. P. Barros-Piedra · E. J. Coronel-González · B. F. Peñafiel-Pinos
Universidad Politécnica Salesiana, Cuenca, Ecuador
e-mail: jordonezo@ups.edu.ec

L. F. Guerrero-Vásquez
e-mail: lguerrero@ups.edu.ec

P. A. Chasi-Pesántez
e-mail: pchasi@ups.edu.ec

D. P. Barros-Piedra
e-mail: dbarros@est.ups.edu.ec

E. J. Coronel-González
e-mail: ecoronel@ups.edu.ec

B. F. Peñafiel-Pinos
e-mail: bpenafiel@est.ups.edu.ec

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_37

1 Introduction

Over the last decade, the rapid evolution of electronic devices has given rise to new technologies and services, most of them based on an Internet connection. Along with this, demand has grown for applications where high transmission speeds are needed for better performance and end-user experience. For example, e-commerce services, social networks, virtualization services, and cloud computing have grown exponentially, and with the advent of the Internet of things (IoT) and fifth-generation (5G) mobile networks, we are forced to restructure traditional network architectures [1].

Most traditional networks were designed based on a hierarchical architecture, which makes sense in a client-server or north-south environment. However, this type of architecture has limitations in the face of new requirements posed by today's technology and is not well suited to the dynamic needs of data centers [2]. The limitations of traditional networks include their limited capacity to adapt to new technologies, low scalability, and inefficient use of access control policies. This has led to the search for alternatives to solve these problems. In view of this, the open networking foundation (ONF) proposes the use of software-defined networks (SDNs) to satisfy current user requirements [3].

SDNs are both an architecture and a design strategy that serve to create a programmable network, also known as east-west networks, in which the control part is decoupled from the hardware part. The control is taken over by a software application called a controller, thus achieving more programmable, automatable, and flexible networks [4]. With these advantages, network administrators gain independence and control over the entire infrastructure from a single logical point, which simplifies design and operation. It also simplifies the use of network devices, as they no longer have to process hundreds of standard protocols, but only have to accept the instructions given from the SDN controllers.

When the network accepts the set of instructions given by the controller, IT operators and administrators optimize their work, since they no longer have to place hundreds of lines of code by hand on N number of devices to achieve a change. This work will be done only at the controller, which in turn will replicate the instruction to the rest of the devices. In addition, by leveraging the centralized intelligence of SDN controllers, the behavior of the network can be altered in real time, and new applications and services can be deployed on the fly.

Currently, the most popular specification for creating an SDN is the open standard called OpenFlow, which was one of the pioneers in realizing a defined communication between the control and data layers. In addition, it allows direct access and manipulation of the data layer of network devices (switches, routers), whether physical or virtual, which means that it is based on a hypervisor. To perform all the work, OpenFlow handles the concept of flow, which allows to identify network traffic when certain predefined rules are met, and based on parameters such as usage patterns, applications, cloud resources, among others, it knows how this flow traffic will pass through the devices.

Applying this technique to wide area networks (WANs), we have SD-WAN technology, which seeks to take advantage of the flexibility and agility of wide area network connections [5]. WAN technology is generally used to connect enterprise networks, which have their data centers and branch offices separated by large geographical distances, this being a limitation for network administrators to modify devices manually. However, with the use of SDN, this problem is overcome, and its advantages are mainly realized. Therefore, SD-WAN solutions offer consistent and pervasive connectivity throughout the network, optimizing application performance, reducing costs, and incorporating agility at all points.

Due to the facilities and advantages of using SD-WAN, the business sector is migrating its networks to this technology very quickly; however, we must be very cautious in terms of cybersecurity. In a traditional corporate network, where services such as MPLS are available, a virtual private network (VPN) is created to transmit data, and this is inaccessible from the Internet. On the other hand, to connect the branches of a company through SD-WAN, it is done directly to the Internet, and in many occasions, they are not secure networks. For this reason, it is necessary that all information transmitted is encrypted, to prevent it from being inspected or edited by third parties. Another problem is that it creates a gateway to the corporate network, and if someone manages to enter, they could alter the entire network. Finally, an additional risk to corporate traffic is the communication mechanism or protocol between the remote devices and the controller [6].

Starting from this context, this paper presents a solution to the security of the controller of an SD-WAN network. This solution is based on the use of blockchain, and for this purpose, a new algorithm known as HUM has been implemented. The main objective is to have several controllers in different points of the network, these can be generated on demand, and in the case that the creation of a new one is required, authorization must be requested to the rest of the controllers, which will give the acceptance, only if the previous creation records are identical, thus providing the centralization of the network in multiple points and the security of the controller in an SD-WAN network.

The paper is organized as follows. In Sect. 2, related works regarding controller safety are presented. Section 3 describes the HUM algorithm. In Sect. 4, the simulation of a new network is presented using the proposed algorithm. In Sect. 5, a possible use case of the proposal applied to a financial institution is shown, leaving for the section 6, the conclusions of the work.

2 Related Works

When talking about security within SD-WAN or SDN controllers, a key point is in the communication channels, since it must be guaranteed that the information is not altered. Therefore, a special concern is the integrity and confidentiality of the data exchanged between the controllers. The use of Firewalls, cloud-based security, IPSec protocols, applications for the use of socket layer security (SSL), and transport layer security (TLS) over the OpenFlow standard are necessary due to the compatibility that SD-WAN networks present in terms of the protocols used by traditional networks.

Although these traditional protocols and applications over SD-WAN offer a higher level of network security, they do not guarantee topology invulnerability, so it is necessary to further analyze the weaknesses of SD-WAN networks, which differ from the traditional security paradigm and fit the needs of the SD-WAN architecture [7].

Thus in [8], a secure and reliable control platform has been developed, in which security issues are specified in seven possible threat vectors: spoofed traffic flows that can be used to attack controllers, attacks on vulnerabilities with the goal of slowing down or breaking communication, attacks on control plane communications by DoS, vulnerabilities in controllers, lack of mechanisms to ensure trust between controller and management applications, vulnerabilities in administrative stations, and finally, the lack of reliable resources for forensic analysis and remediation.

Security in enterprise networks differs from that of the Internet. Thus, the secure architecture for the enterprise network or SANE presented in [9] is a proposal in which networks can be managed through a centralized and authenticated control by all network elements to ensure the security of the enterprise through simple and strong high-level policies, which are independent of the topology and the equipment used. Keeping the application at the link layer to prevent lower layers from weakening it, it also hides information about the topology and services from those who do not have permission to see them, thus maintaining a single central trust component, where all policy is defined and executed centrally.

Also, at [10], they present Ethane, a system that further enhances the SANE architecture using a centralized controller. Ethane manages routing, flow admission, and couples simple Ethernet switches based on the required flow. One of Ethane's most powerful features is that it names devices so that it can easily track all links between names, addresses, and physical ports on the network, for which it authenticates between all switches using multiple methods. Ethane and SANE are designed to enable secure communication between the control plane and the data plane.

In [11], we can find an application called OpenSec, which is based on the OpenFlow driver, it allows network operators to describe security policies using human-readable language and implement them throughout the network. OpenSec acts as a virtual layer between the user and the complexity of the OpenFlow controller and automatically converts security policies into a set of rules that are entered into network devices.

In the case of defense against distributed denial of service (DDoS) attacks, [12] introduces an autonomous defender based on OpenFlow-enabled switch combining OpenFlow and Locator/ID separation protocol (LISP) technologies. The experiment emulates 100 attackers to a server, sending a total of 1000 packets per second (PPS) to the server, where the DDoS defender monitors the OpenFlow switch flows and detects the DDoS attack through volume counting. The defender's design is based on a closed loop where the first threshold is 3000 packets per 5 seconds. Once the traffic value exceeds the threshold, second stage detection is triggered. If the traffic reaches 800 packets per second 5 times in a row, the DDoS defender drops the incoming packets. After the flow entry times out, the defender returns to normal and will control the traffic volume.

Finally, in [13], they demonstrate how TLS support over OpenFlow has an impact on packet input delays from SDN switches due to the notable impact, it has on the driver, so it is necessary to add hardware acceleration support for encryption to future OpenFlow switches. For an SDN deployment in productive networks, encryption is unavoidable, but at present, especially, the control plane and test software are not widely supported. In this way, in this paper, they apply the use of Open vSwitch software, which has lower packet input delay compared to TLS over OpenFlow applications.

As can be seen, there are different ways to secure software-defined networks and the SD-WAN variant, some of them clearly centered on the controller, or on the communication system between the control plane and the data plane. However, there is no alternative where the controller is intended to be decentralized at multiple points. For this reason, the HUM algorithm is presented in the following section.

3 HUM Algorithm

HUM, a blockchain-based algorithm for concurrent OpenFlow controllers, is proposed. This algorithm consists of a series of controller nodes, which are connected to the network at various points. These nodes can be generated on demand, for example, using virtualized environments, thus contributing to the scalability and availability of the network.

Each controller runs the algorithm independently and maintains a copy of the network state in its own permanent storage. This copy is stored as a chain of blocks connected by the information in each block. A block consists of the following fields:

- **Block identifier**. It is a hash code that allows to determine the identity of the block to chain it with others and additionally allows to verify that it has not been modified. It is built using blake512 on the concatenation of the other fields.
- **Counter**. It is a positive integer sequential number that indicates the position of the block in the chain and is additionally used for the message consensus system.
- **Timestamp**. A 64-bit integer value in TAI64 format indicating when the block was created. It is base64 encoded.
- **Creator**. It is a public key curve25519, of 32 bytes belonging to the author of the block, it is encoded in base64.
- **Content**. It is the text in Json format, which contains the network configuration information, be it configuration commands for controllers, static openflow flow tables for switches, or programs for handling reactive packets.

When the entire history of the flow changes is found in the blockchain, synchronization between the nodes is maintained, and a new node or a node that has recovered from a crash will regain functionality by simply requesting the missing blocks in its registry from the rest of the nodes in the net.

Keeping a record of all the operations carried out is very useful in case of security audits or to return the network to the previous state in case of problems with a new configuration. HUM assumes a proactive SDN model in which flows are imposed by drivers, and packets that arrive at an OpenFlow device without a matching rule will simply be dropped. All controllers have a total view of the network, and different devices can be connected to any of the controllers to update the data streams, allowing load balancing on the controller.

Each node is identified by a public key, the impersonation of the nodes should only be possible if they seize the private key of any of them, and in case of entering the network, it would be registered as part of the events within the blockchain and by the nature of it is virtually impossible to alter to hide the trail making it difficult to maintain anonymity.

For the generation of hashes that help with the referencing of the nodes, blake512 will be used. This system does not use TLS but is based on WireGuard, which is a protocol where each node is identified by a 32-byte public key. To establish communication with each other, they only need a round trip, this protocol is silent, so there is no way to probe the daemon to know if it is listening. In addition, a time stamp is placed on the exchanged messages, to avoid replay attacks.

One of the nodes of the network at any time can add a new block to the chain, for this, it sends the block to the rest of the nodes as a distributed transaction. In case, the transaction is accepted by the majority of the nodes; here, the majority is made up of half plus one of the nodes, and the transaction is considered successful and the nodes proceed to add the new block to their own chains in permanent storage. The reason why a majority is necessary is because the consensus algorithm used so that the nodes can decide whether to accept or reject the creation of a new block is paxos. Paxos tells us that given a network of N nodes it only needs the presence of a majority number of nodes that commit to accept a transaction to ensure consistency. Paxos has a series of requirements to function, the main one being that when a node commits to a transaction it must be able to remember the commitment and that the transactions have to be numbered in ascending order. Both requirements are met thanks to the chain of blocks that permanently maintains in the record of all transactions (blocks) accepted.

When a node starts up for the first time, it proceeds to generate the initial block, also called the genesis block or block 0. Then, an existing node, if any, makes the request to register it on the network. In addition, the new block is registered in the chain, a step that is indicated to the rest of the nodes. If the new node is the first in the network, it assumes its own registration.

For the synchronization of the nodes, the consensus process requires only the majority of the nodes, some nodes may be temporarily out of date with the information of the blockchain, so they periodically consult the rest of the nodes about the existence of new blocks. If found, they are automatically added to the end of the chain.

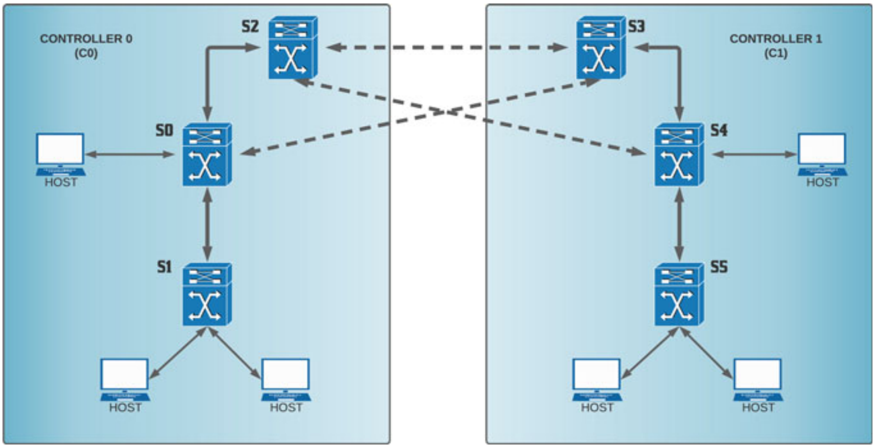


Fig. 1 Simulated network topology

4 Simulation of a New Network

This section presents how the HUM algorithm was built and simulated, for which Node.js was used, and simulations were performed using OpenVswitch. In Fig. 1, you can see the simulated topology, it should be noted that when using OpenVswitch to perform the simulations, and we will call the forwarding devices as a switch. In this case, all switches are capable of contacting existing controllers, although, ideally, there is no need. Only when a switch fails to contact its controller, it proceeds to try to receive the configuration of the rest of the existing controllers on the network.

For the simulation process, the aforementioned network was built in a virtual machine, after that, the services in the controllers were initialized using the flow indicated in Fig. 2.

The process shown is detailed below:

0. Controller 0 starts up, and when it is empty, it proceeds to create a genesis block and automatically save it in the chain. Additionally, the identity of the controller (public key) and the private key are generated.
1. Controller 0 is instructed to register itself, thus becoming the first node on the network.

2a. Controller 0 is instructed to register the flows to be configured. In the controller, it creates a configuration block, and once stored, it proceeds to distribute the configuration rules among the devices belonging to the network. At this time, being the only node, it automatically becomes the controller of the entire network.

2b. At the same instant of time, controller 1 starts and also generates the genesis block, its identity, and its private key.

3. Controller 1 asks controller 0 to add it to the consensus network. Controller 0 accepts and registers the block with the information from controller 1.

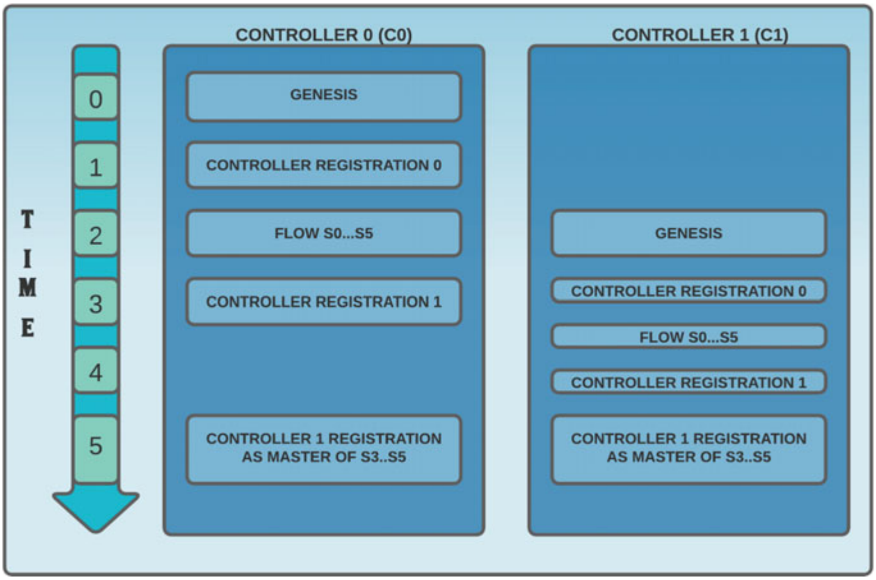


Fig. 2 Blockchain status on controllers 0 and 1

4. Controller 1, which is now part of the consensus network, proceeds to fill in the missing information; for this, it makes the request for the missing blocks to controller 0. It should be noted at this point that the failure of any of the controllers would paralyze the consensus network, since by themselves, none of the controllers has a majority. This is in contrast to the pre-registration state of controller 1 where controller 0 had an absolute majority and could freely add blocks to the chain.
5. Controller 1 wishes to become the primary controller for switches s3, s4, and s5. For this, it creates the new block and asks the consensus network, which now includes it, to accept this change. Once approved, both nodes have the configuration block added to their chains simultaneously, none of them needed to be matched after the admission of new blocks.

Once the network is working, any of the controllers can be removed, as an example we place the controller 0, and the switch devices will be able to maintain their configuration by contacting the alternative node. However, the consensus network would be frozen, for this reason, new changes could not be introduced at this time, since it is impossible to get a majority. You cannot introduce new nodes to the network, since that also requires consensus, in the same way, you cannot remove controller 0 from the network registry because this also requires consensus. If controller 0 does not return, the network is effectively frozen forever. This was done in the simulation, and it was seen that they actually followed the expected guidelines. However, in the case that we have three controllers c0, c1, and c2 in the consensus network, and c0 is permanently lost, then we can simply create a new node c3, register it in the network (c1 and c2 would be the majority) and register the removal of node c0 from the network.

A node being permanently lost would mean that a node cannot be enabled with the original public/private keys. If the keys of a backed-up node are kept, the recovery of the node would only consist of booting a new node with these keys. The node would proceed to equalize again requesting information from the blocks of the rest of the network and once equalized it would be fully operational.

5 Use Case: Application of the Proposal in the Business Network of a Financial Institution

Finally, in this section, a possible use case of the proposed algorithm is detailed. Given the case of a financial institution, our proposal would make it possible to increase security, an extremely important and at the same time critical element in companies with this line of business. An example will be given to a financial institution which has a headquarters located in a city, and different branches distributed throughout the country, all of them geographically very distant. All branches of the entity are connected by means of a private SD-WAN network; therefore, the integrity of the information that circulates in the network is essential for the good performance of the business. Normally, a central controller would be placed in the institution's headquarters, the same one that would be in charge of maintaining the configuration of the entire network, and if necessary, it would make changes. However, it could be the case that a branch is disconnected for any factor, in this example, network services only of that branch would be completely lost. But there could be a more critical case where the controller is disconnected from the matrix, either due to a system failure, human error, or an intentional attack, because of this the entire network would collapse. Given this, a possible solution is to place multiple controllers one in each branch, whose function would be to maintain not only the flows of its own branch, but also to maintain the flows of the entire financial institution. Thus presenting a solution to keep the network operational at all times, counting on the ubiquity of the controllers. This solution can be seen in Fig. 3, where the network of a financial institution is shown, and each branch has an SD-WAN controller.

6 Conclusions

The rapid growth of electronic devices has meant that traditional networks have had to migrate to networks that provide greater capabilities and benefits. For this reason, today, there are software-defined networks, which provide better features and qualities to meet the technological challenges that arise. One application of

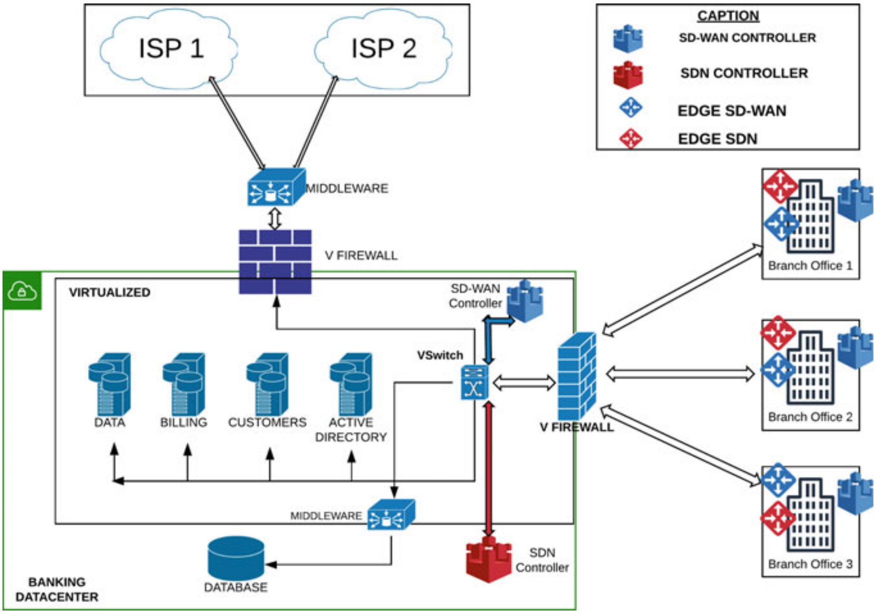


Fig. 3 Solution using HUM algorithm in a financial institution

this type of networks is those focused on wide area enterprise networks or also known as SD-WAN, and these networks cover the needs of companies that have different branches in large geographical areas. However, due to their dependence on a central controller, they have inherent security problems, especially in cases where they can communicate with public networks such as the Internet. The distributed algorithm based on a blockchain architecture aims to eliminate or reduce some of these problems. Finally, the use case presented shows a viable implementation of the algorithm in a financial entity that could be implemented with few modifications in any entity with similar requirements to distribute its systems over WANs.

References

1. Gallegos-Segovia PL, Bravo-Torres JF, Vintimilla-Tapia PE, Ordoñez-Ordoñez JO, Mora-Huiracocha RE, Larios-Rosillo VM (2017) Evaluation of an SDN-WAN controller applied to services hosted in the cloud. IEEE Second Ecuador Technical Chapters Meeting (ETCM). IEEE, pp 1–6
2. Spera C (2013) Software defined network: el futuro de las arquitecturas de red. Logicalis Now, pp 42–45
3. Chico JC, Mejía D, Bernal I (2013) Implementación de un prototipo de una red definida por software (SDN) empleando una solución basada en hardware. Escuela Politécnica Nacional, Quito, Ecuador

4. Figuerola N (2013) Sdn-redes definidas por software. línea]. Disponible en: https://articulositfiles.wordpress.com/2013/10/sdn.pdf
5. Wang DW (2018) Software defined-WAN for the digital age: a bold transition to next generation networking. CRC Press
6. Tejedor E (2016) Retos de seguridad en las nuevas redes sd-wan. Seguritecnia 434:111–113
7. Jain R, Khondoker R (2018) Security analysis of SDN-WAN applications-b4 and iwan. In: SDN and NFV security. Springer, pp 111–127
8. Kreutz D, Ramos F, Verissimo P (2013) Towards secure and dependable software-defined networks. In: Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking, ACM, pp 55–60
9. Casado M, Garfinkel T, Akella A, Freedman MJ, Boneh D, McKeown N, Shenker S (2006) Sane: a protection architecture for enterprise networks. In: USENIX security symposium, vol 49, p 50
10. Casado M, Freedman MJ, Pettit J, Luo J, McKeown N, Shenker S (2007) Ethane: taking control of the enterprise. ACM SIGCOMM Comp Commun Rev 37(4):1–12
11. Lara A, Ramamurthy B (2014) Opensec: a framework for implementing security policies using openflow. In: IEEE global communications conference. IEEE, pp 781–786
12. YuHunag C, MinChi T, YaoTing C, YuChieh C, YanRen C (2010) A novel design for future on-demand service and security. In: 2010 IEEE 12th international conference on communication technology. IEEE, pp 385–388
13. Durner R, Kellerer W (2015) The cost of security in the SDN control plane. In: ACM CoNEXT 2015-student workshop

Hybrid Methods to Analyze a Skin Tumor Image and Classification

Asmaa Abdul-Razzaq Al-Qaisi and Loay E. George

Abstract Processing medical images involve the creation of problem-specific strategies for improving raw medical imaging data for targeted visualization objectives and additional research. There are numerous medical issues, some place emphasis on broadly applicable theories, and some concentrate on certain uses. We mainly concentrate on segmenting images and doing multi-spectral analysis. After the image is preprocessed and the tumor area is isolated, a hybrid between two methods, sub-block discrete cosine transform and second-level discrete wavelet transform, is used to transform the image into a frequency domain to analyze the tumor area and calculate the features; features are computed in two ways; the first way is Michelson contrast (calculated after the image is preprocessed), and the second way is first-order statistics that calculated after the sub-block DCT and two-level DWT performed, the output saved as a feature vector after that passes the vector to backpropagation NN. For training and classification, dataset ISIC 2018 was used in the experimental analysis (we use only four cases). ANN was used for classification, and the results show that it is accurate to roughly 88.98% for DWT, and 85.44% for sub-band DCT, the ANN training performance (MSE) after 1000 epochs for first-order statistics of (DWT + DCT + Contrast) is $2.69 * 10^{-4}$.

Keywords Color image · Skin tumors · Classification · Feature extraction · DCT · DWT · Statistical methods

A. A.-R. Al-Qaisi (✉)
College of Education for Women, Baghdad University, Baghdad, Iraq
e-mail: asma_72@coeduw.uobaghdad.edu.iq

L. E. George
University of Information Technology and Communication, Baghdad, Iraq
e-mail: loayedwar57@uoitc.edu.iq

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_38

1 Introduction

The most prevalent cancer in the world, skin cancer, is noted for its increasing prevalence and increasing burden. It can be challenging to visually discern between normal and abnormal tumors with general malignancies [1].

The unchecked growth of abnormal cells in the skin is recognized as skin cancer. It takes place whenever skin cells experience DNA damage resulting in modifications, or genetic flaws, that causes the skin cells to proliferate swiftly and develop dangerous malignancies. Physical examination and surgery are typically used to diagnose skin cancers [2]. The surgery is a straightforward procedure in which all or part of the area is removed and tested in the testing facility. Doctor's experiment based on eye extracts features from the data. An ANN has been used. The most effective procedure is dermoscopy to produce color imaging of skin, which equipment has made the most advancements in the research of cancers [3]. The common skin cancer datasets are small and only include a few different forms of the disease along with a small number of photos for each [4].

Data and features are extracted. The ANN was utilized. Researchers' efforts to develop ideal cancer classification systems include developing methods for classifying skin tumors. The convolutional neural network (CNN) presentation works on using skin tumor image data. Duplicated data and features are extracted. The ANN was utilized, by rotating and flipping in [5] data, and features are extracted. The ANN was utilized and has a favorable effect. It corrects for differences between the test and training sets. NN was employed in [6] to divide skin cancer into three categories. This technique uses a particular sort of dermoscopic picture that operates directly on color skin image without preprocessing. In [7], the additional dense NN and CNN are used to balance the characteristics. This effort explains the methods of the dataset by adding outside data with a category for skin tumors. The extreme class imbalance in the number of photos is the other class lesion issue. Input resolutions for the die rent method and die rent cropping strategies are considered in this study, combining data based on factors such as age, structural location, sex, etc. to take care of this property [8]. The classification of three-dimensional images is by converting them into binary images. Utilizing the segment characteristics and ANN classification, the novel algorithm known as Adaptive Snake has been used.

The purpose of the ANN model, which employs many dispensation layers, is to provide accurate abstractions of data. The following are the key components of using ANN: (i) the reduction of large amounts of cancer-related data used to train NN models; (ii) the development of graphics processing units computer analysis; and (iii) data and features are extracted. In ANN approaches, the ANN has been utilized for batch normalizing, dropout correction, and linear unit correction [9].

Extract features from the data. An ANN has been used. The automatic tumor identification utilizing color dermoscopy images exhibits three intriguing agreements. First, the size, texture, color, and shape of the lesions on the skin are very similar to those that are typical of many courses. The second reason is the strong association

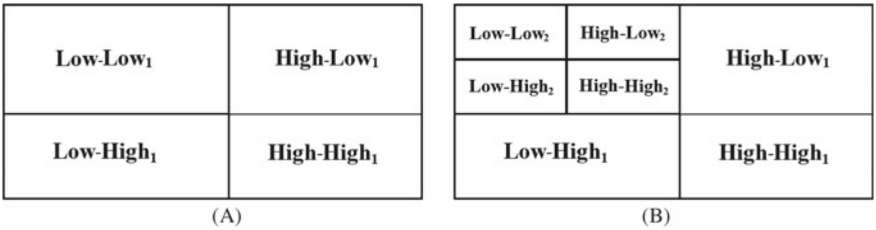


Fig. 1 DWT sub-band, **a** 1-level DWT and **b** 2-level DWT [14]

between lesions with and without melanoma. Thirdly, a variety of environmental elements include hair, noise, veins, and brightness [10–12].

In this study, two types of features with first-order statistic computation for preprocessing are used. Two cascade feature extraction techniques have been applied in this paper. First, the (DCT) and (DWT) transformer methods have been applied. To reduce the data and extract features, statistical methods that determine mean, standard deviation, skewness, and kurtosis have been implemented. Types of cancer have been identified and categorized using the ANN.

2 Discrete Wavelet Transform Technique (DWT)

Data and features are extracted. The ANN was utilized. The wavelet transforms change the picture pixels' spatial and frequency dimensionality (DWT coefficients). Filter bands, a collection of low-pass and high-pass filters, are included. To generate multi-level DWT, a filter may be used in DWT. To eliminate the pixels' split into numerous frequency bands, the DWT employs filter bands. To gather, the DWT converts the pixels into multi-scale representations of both the spatial and frequency self. These recommendations are for efficient multi-scale exploration with less expensive calculations [13]. It generates location-sensitive data that is vital to understanding thyroid nodules. Figure 1 displays the sub-bands for 1-level and 2-level wavelet decomposition utilizing two-dimensional DWT Haar filters.

3 Discrete Cosine Transform (DCT)

The DCT is the actual component of the discrete Fourier transform, which is a mathematical equation that converts signals to the frequency domain (DFT). When the frequency increased in a zigzag pattern, it moved from the left top corner to the right bottom corner [15, 16]. Each pixel in a size of picture of ($N * M$) is indicated by the letter p (x, y). The image is converted into the P (u, v) DCT coefficient using the two-dimensional DCT:

$$P(u,v) = \frac{1}{\sqrt{M+N}} C(u)C(v) \sum_{x=0}^{N} \sum_{y=0}^{M} p(i,j) \cos\left(\frac{(2x+1)2\pi}{2N}\right) \cos\left(\frac{(2y+1)2\pi}{2M}\right) \tag{1a}$$

$$C(u), C(v) = \left[\begin{array}{ll} 1/\sqrt{2} & u, v = 0 \\ 1 & u, v \neq 0 \end{array}\right. \tag{1b}$$

4 The First-Order Statistical Statistics

Many statistical techniques have been applied to skin cancer photos in order to extract data features [17]. Color denotes variance skewness and forward motion, among other things. The statistic equation utilized in this work is fairly straightforward and invariant to information about an image's transformation coefficients [18–20].

I. **Mean**: The mean of coefficients $X(i, j)$ of size M, N is

$$\text{Mean} = \frac{1}{N \times M} \sum_{x=1}^{N} \sum_{y=1}^{M} p(x, v) \tag{2}$$

II. **Standard Deviation**

$$\text{Std} = \sqrt{\frac{1}{N \times M} \sum_{x=1}^{N} (p(x, y) - \text{mean})^2} \tag{3}$$

III. **Skewness**

$$\text{Skewness} = \frac{1}{N \times M} \sum_{x=1}^{N} (p(x, y) - \text{mean})^3 \tag{4}$$

IV. **Fourth Momentum (kurtosis)**

$$\text{Fourth Momentum} = \frac{1}{N \times M} \sum_{x=1}^{N} (p(x, y) - \text{mean})^4 \tag{5}$$

V. **Contrast** [21]: The segmented sub-region of size ($n*n$) has contrast by

$$\text{Contrast} = \sum_{x=1}^{n} \sum_{y=1}^{n} \frac{|p(x, y) - \text{mean}_{n*n}|}{\text{Std}_{n*n}} \tag{6}$$

VI. **Michelson Contrast**

$$\frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{7}$$

5 ANN

Three layers make up the ANN: the input and output with one or more hidden layers [21]. The values of the characteristics, such as DCT, DWT, or statistical approaches (used in this work), are dependent on the training methods to create the networks accurately [22]. In order to reduce the discrepancy between calculated output values and expected values, the network is constructed using computed layer weight values. The values are updated by repeated calculation. The vanishing gradient is the name given to this issue [23, 24].

6 Skin Tumor Classification Methods

Various methods for classifying skin images have been used in this work. Figure 2 displays the suggested systems block diagram.

The systems stages

I. **Image preprocess**

The preprocessing processes to separate the cancerous component from another skin part are as follows (algorithm A):

- Noise removal: In this step, median filters are used to remove extraneous pixels from RGB photographs, such as hair.
- Cropping: The cancer zone is our focus of interest. Cancer's bordering white spaces have been trimmed.
- Thinning: By removing a few foreground pixels.

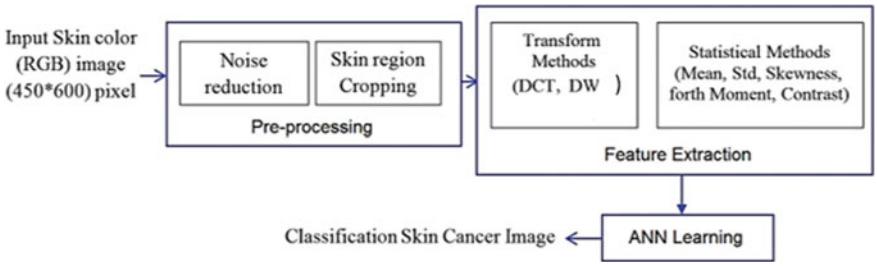


Fig. 2 Skin tumors' classification systems

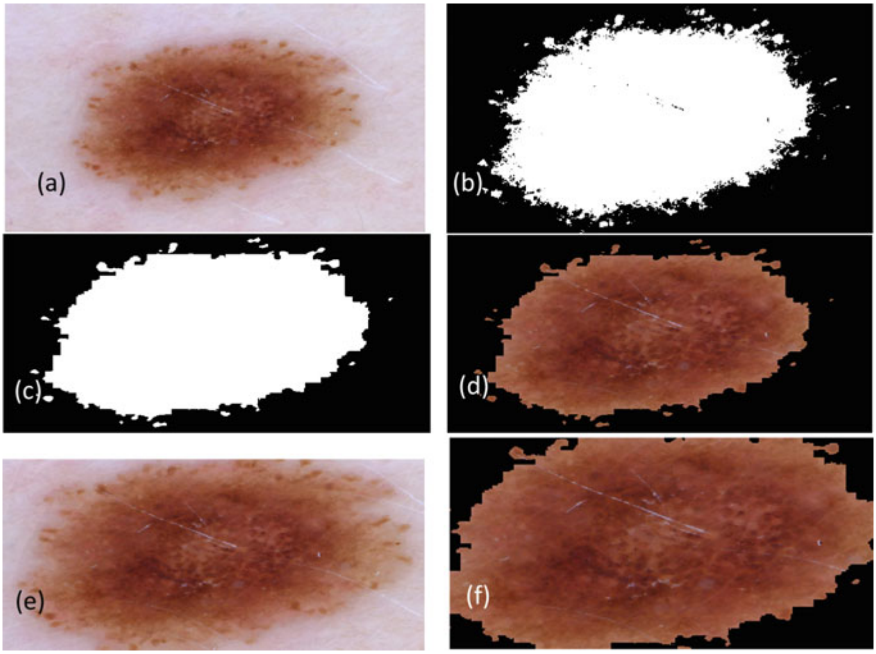


Fig. 3 Preprocessing system, **a** original image, **b** median filter for image, **c** thinning the cancer area, **d** sounding cancer region, **e** normalized the image, **f** normalized cancer region to 128 × 128 pixel

- Normalization: Only the cancer patches are cropped from the three 128 × 128-pixel RGB pictures. Figure 3 shows the overall preprocessing stage.

II. **Feature Extraction**

The feature extraction of the proposed system has two steps and shows in algorithm (B). The normalized cropped cancer region has been converted into frequency domain as the initial stage. To extract and transform the RGB image's pixel differences, two alternative transformations (DCT, DWT, as previously explained) were applied.

- To extract features from the surrounding area, the suggested system separated the cropped image into sub-bands (8 * 8), with each band having a size of (16 * 16). This is seen in Fig. 4.

Figure 5 shows the DWT coefficient for color image, 1-level and 2-level.

Second Step: First, frequency domain was applied to the normalized cropped cancer region. To extract the difference between each image pixel and transform it, two

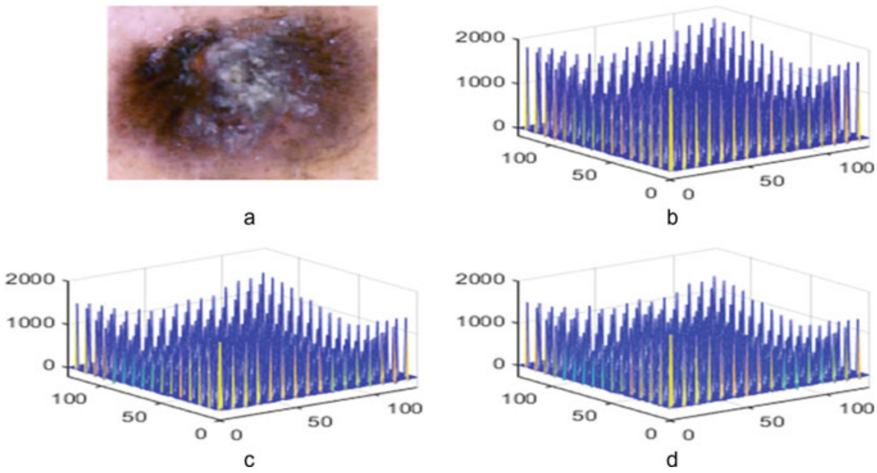


Fig. 4 Sub-block DCT for normalized cancer image, **a** original image, **b** coefficient of R image, **c** coefficient of G image, **d** coefficient of B image

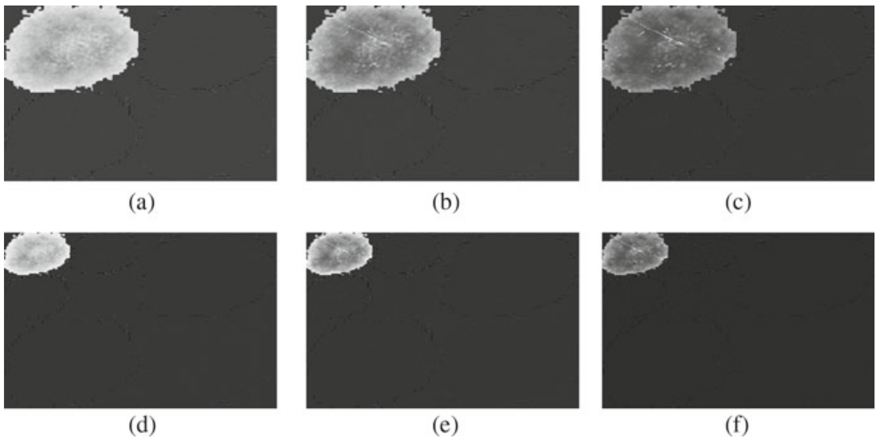


Fig. 5 DWT for normalized cancer image, **a**–**c** the 1-level DWT coefficients of R, G, and B image, respectively, **d**–**f** the 2-level DWT of R, G, and B image, respectively

distinct transformations (DCT, DWT, as explained above) were applied to the RGB data.

- The suggested system segmented the cropped image into sub-bands (8*8), with each band having a size of (16*16), and implemented the DCT for each sub-band in RGB images to extract characteristics from the neighborhood; feature vector illustrated in Figs. 6 and 7 shows the diagram of hybrid method.

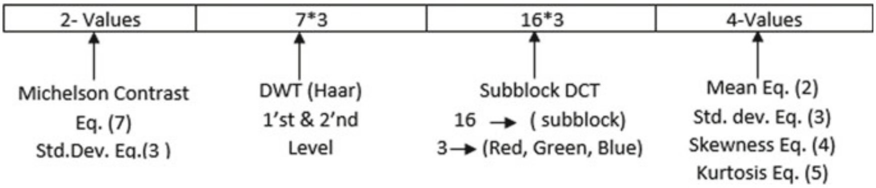


Fig. 6 Hybrid features vector

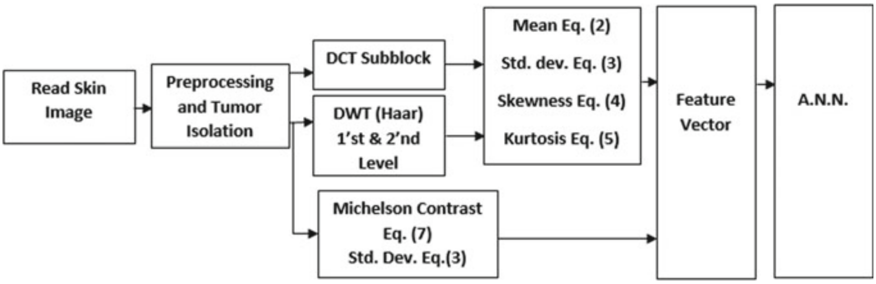


Fig. 7 Diagram of hybrid method

III. ANN

An artificial network has been constructed using the backpropagation technique for training and classifying data related to skin cancer. The ANN has two hidden layers, each of which is made up of 20 nodes. The error has been reduced with the 200 iterations.

Algorithm (A): Preprocessing Image

Input: File Image of Skin cancer
Type of Data: Image (JPEG)
Size of Image: 450x600

Output: Color Image Preparation.
Image Segmentation.
Size reduction

Begin
Step 1: perform Segmentation with Resize Image
List 1 ← **Image Reading (450x600)**
List 2 ← **Binarization**
Step 2: Isolate area of interest
List 1 ← Reading **Binary Image**
List 2 ← **(5x5) Average Filter**
List 3 ← **(17x17) Full Space Average Filter**
List 4 ← **(128*128) Resize**
End

Algorithm (B): Hybrid (Contrast+ DWT+DCT)
Input: Preprocessing Image **Type of Data** : JPEG Image **Image Size**: 128x128
Output: Feature Vector **(7x3x4)+(16 x3x4)+(2*3)=282**
Begin **Step 1: Contrast ( )** Michelson contrast Eq. (7) **Std.Dev. Eq. (3)** **Step 2: First Level of DWT** **L1 ⟵ Read image** **L2 ⟵ DWT 4x3** **Step 3: Second Level of DWT** **L1 ⟵ Read image** **L2 ⟵ DWT 7x3** **Step 4: Subblock DCT** **List1 ⟵ Read Skin Color image** **For i (where 0 ≤ i ≤ 128)** **For j (where 0 ≤ j ≤ 128)** **ii= i div 4** **jj=j div 4** **calculate DCT ()** **end** **end** **Step 5: 1'st order statistic** **List3 ⟵ Mean Eq. (2)** **List4 ⟵ Std. Dev. Eq. (3)** **List5 ⟵ Skewness Eq. (4)** **List6 ⟵ Kurtosis Eq. (5)** End

7 Output of the Systems

The seven types of tumor images has been used (Table 1), and 70% of the dataset has been used to train and 30% used for testing the classification methods.

The proposed hybrid method gives feature extraction from sub-band DCT; the one- and two-level DWT with the second step four statistic has been calculated to extract the minimum size of features for that the number of features extracted in DCT is (color Image have $4 * 3 * 4 = 48$ feature), with DWT are (color Image have $11 * 3 * 4 = 123$ feature). The proposed ANN structure consists of two hidden layers (20 nodes for every layer) and one output layer with epoch = 1000 with performance 2.69×10^{-4} as shown in Fig. 8.

Table 1 Numbers of tumor images in dataset

No.	Tumors type	No. of image
1	Benign (B) leratosis (KL)	1099
2	Dermatofibroma (DF)	115
3	Vascular (VASC) lesions	142
4	Melanoma (MEL)	1113
5	Nevus (NV)	6705
6	Basal (B) cell (C) carcinoma (C) (BCC)	514
7	Actinic (A) keratosis (KIEC)	327

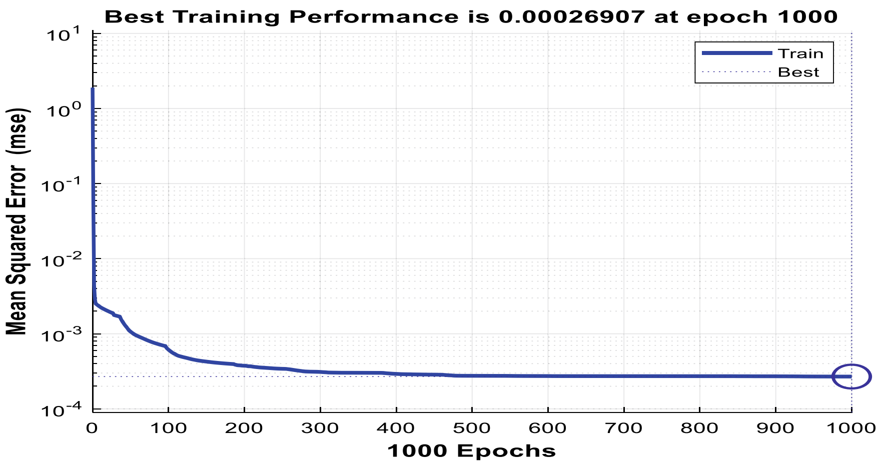


Fig. 8 Train performance for DCT + DWT (hybrid) feature extraction for 1000 epoch, the MSE = 2.69 * 10^{-4}

The confusing matrix has been calculated for the skin tumor image implemented as training and testing features for ANN systems. Table 2 shows the confusion matrix for DCT, and Table 3 shows the confusion matrix for DWT, and the results of testing ANNs of the DCT + DWT + Contrast feature extraction method have been implemented in Table 4 to classify the seven tumors' types, and Table 5 shows the MSE.

8 Conclusions

Analyzing the dataset reveals inequity, with a big disparity in the total number of photos for each category, making it difficult to simply categorize image attributes for pests. The outcome displayed in Table 2 demonstrates that DWT + DCT + Contrast approaches provide a high level of accuracy in the classification of cancer. The danger

Table 2 Confusing matrix for DCT + contrast method

Actual tumor types	Tumor type	Predicted tumor types						
		1	2	3	4	5	6	7
	1	242	72	9	4	0	0	0
	2	23	414	67	10	0	0	0
	3	2	65	767	252	11	2	0
	4	0	0	15	89	8	3	0
	5	0	0	98	380	5982	240	5
	6	0	0	30	123	203	702	55
	7	0	0	0	0	0	15	125

Table 3 Confusing matrix for DWT + contrast method

Actual tumor types	Tumor types	Predicted tumor types						
		1	2	3	4	5	6	7
	1	212	104	11	0	0	0	0
	2	42	397	74	1	0	0	0
	3	1	96	892	108	2	0	0
	4	0	0	17	93	3	2	0
	5	0	0	105	342	6125	128	5
	6	0	0	14	39	110	922	28
	7	0	0	0	0	1	8	131

Table 4 Confusing matrix for DCT + DWT + contrast (hybrid method)

Actual tumor types	Tumor types	Predicted tumor types						
		1	2	3	4	5	6	7
	1	308	19	0	0	0	0	0
	2	9	492	13	0	0	0	0
	3	0	30	1014	50	5	0	0
	4	0	0	10	104	1	0	0
	5	0	0	74	212	6324	95	0
	6	0	0	0	3	50	1042	18
	7	0	0	0	0	0	0	140

		Predicted tumor type	
		Positive	Negative
Actual tumor type	Positive	6883	140
	Negative	451	2541

Table 5 ANN training performance (MSE) after 1000 epoch

No.	Proposed hybrid method	MSE
1	1st statistic of (DCT) + contrast methods	7.98×10^{-4}
2	1st statistic of (DWT) + contrast methods	7.41×0^{-4}
3	1st statistic of (DCT + DWT) + contrast methods (hybrid-3)	2.69×10^{-4}

is reduced, and the method's robustness is raised when the DCT subdomain provides a lower error detection rate than the other types.

The classification of tumors types by using this method gives higher identification and reaches to 100% classification for type seven, 94% for type one, and 90.4% for type two with performance reach to 2.69×10^{-4} as shown in Fig. 8.

9 Future Prediction

There are other assessment tools for skin tumor image that can be used besides those used in the study, which is the use of co-occurrence matrix with second-order statistics and its combination with the mentioned methods to research in addition to other types of classification to find a new and more accurate pattern in diagnosis.

References

1. Alam MA, Autonomous M (2020) Automated skin lesion classification using ensemble of deep neural networks in ISIC 2018: skin lesion analysis towards melanoma detection challenge. University of Barcelona Campus de la UAB, Plaça Cívica, 08193 Bellaterra, Barcelona, Spain
2. Satheesha TY, Satyanarayana D, Giriprasad MN, Nagesh KN (2016) Detection of melanoma using distinct features. In: 3rd MEC international conference on big data and smart city
3. Hosny KM, Kassem MA, Fouad MM (2020) Skin melanoma classification using deep convolutional neural networks. In: Deep learning for computer vision: theories and application. CRC Press, Boca Raton, FL, USA
4. Han SS, Kim MS, Lim W, Park GH, Park I, Chang SE (2018) Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. Published by Elsevier, Inc. on behalf of the Society for Investigative Dermatology, pp 1529–1538
5. Codella NCF, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, Kalloo A, Liopyris K, Mishra N, Kittler H, Halpern A (2017) Skin lesion analysis toward melanoma detection. In: International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). arXiv:1710.05006 [Online]. Available: http://arxiv.org/abs/1710.05006
6. ShiyamSundar RS, Vadivel M (2016) Performance analysis of melanoma early detection using skin lesson classification system. In: International conference on circuit, power and computing technologies (ICCPCT)
7. Gessert N, Nielsen M, Shaikh M, Werner R, Schlaefer A (2019) Skin lesion classification using ensembles of multi-resolution EcientNets with MetaData. arXiv:1910.03910v1[cs.CV]

8. Mohan Kumar S, Ram Kumar J, Gopalakrishnan K (2019) Skin cancer diagnostic using machine learning techniques, wavelet transform and naïve Bayes classifier. Int J Eng Adv Technol (IJEAT) 9(2):2249–8958
9. MATLAB Central Program or Color Image Segmentation—Ath Narayan. College of Engineering, India, 15 Aug 2018. https://www.mathworks.com/matlabcentral/fileexchange/25257-color-image-segmentation?focused=5191437&tab=function
10. Ha H, Man T (2018) Against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 29(8):1836e42
11. Barata C, Celebi ME, Marques JS (2019) A survey of feature extraction in dermoscopy image analysis of skin cancer. IEEE J Biomed Health inform 23(3)
12. Monika MK, Vignesh NA, Kumari ChU, Kumar MNVSS, Laxmi E (2020) Lydia Skin cancer detection and classification using machine learning materials today. Elsevier, Proceedings journal homepage: www.elsevier.com/locate/matpr
13. Seal A, Bhattacharjee D, Nasipuri M (2017) Predictive and probabilistic model for cancer detection using computer tomography images. Multimed Tools Appl 77:3991–4010
14. Brinker TJ, Hekler A, Utikal JS, Grabe N, Schadendorf D, Klode J, Berking C, Steeb T, Enk AH, von Kalle (2018) Skin cancer classification using convolutional neural networks: systematic review. J Med Internet Res 20(10):e11936
15. Nahata H, Singh SP (2020) Deep learning solutions for skin cancer detection and diagnosis. https://doi.org/10.1007/978-3-030-40850-3_8
16. Almeida AM, Santos IAX (2020) Classification models for skin tumor detection using texture analysis in medical images Marcos. J Image 6:51
17. Tschandl P, Rosendahl C, Kittler H (2018) The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 5:180161
18. Combalia M, Codella NCF, Rotemberg V, Helba B, Vilaplana V, Reiter O, Carrera C, Barreiro A, Halpern AC, Puig S, Malvehy J (2019) BCN20000: Dermoscopic lesions in the wild. arXiv:1908.02288. Available: http://arxiv.org/abs/1908.02288
19. Haralick RM (1979) Statistical and structural approaches to texture. Proc IEEE 67:786–804
20. Bahadure N, Ray AK, Thethi HP (2017) Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. Int J Biomed Imaging 2017:1–12
21. Abdel-Nasser M, Moreno A, Puig D (2019) Breast cancer detection in thermal infrared images using representation learning and texture analysis methods. Electron 8:100
22. Ayyachamy S (2015) Registration based retrieval using texture measures. Appl Med Inform 37:1–10
23. International Skin Imaging Collaboration. Available online: https://challenge2019.isic-archive.com/. Accessed on 2 Dec 2019
24. Brinker TJ, Hekler A, Enk AH, Klode J, Hauschild A, Berking C, Schilling B, Haferkamp S, Schadendorf D, Holland-Letz T (2019) Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. Eur J Cancer 113:47–54

Funnel Control for Multi-agent Systems in a Disconnected Condition

Hiroki Kimura and Atsushi Okuyama

Abstract A multi-agent system (MAS) consists of multiple autonomous agents. A figure composed of agents and the connections between them is called a graph. The overall behavior of the MAS is determined by the local interactions among its agents. MAS is based on graph theory, and recently attention has been focused on approaches to control theory that considers network structures. A control method that converges the states of all agents is called consensus control. In this study, the agents are assumed to be actual robots, and their communication range is assumed to be finite. A graph may be disconnected when the communication range of the agent is limited. Therefore, we studied the consensus problem of MAS by considering the effects of these disconnected conditions. We applied funnel control, which considers disconnected conditions as a control method to achieve consensus. Funnel control is a high-gain adaptive control method that can guarantee tracking with a preset degree of accuracy and suppresses deviation within a predefined function. However, this function is not uniquely obtained. We performed a simulation study to demonstrate the effectiveness of the proposed method for solving the MAS consensus problem.

Keywords Multi-agent systems · Disconnected condition · Funnel control

1 Introduction

A multi-agent system (MAS) is comprised of multiple agents, and the behavior of the entire system is determined by the local interactions among the agents [1]. In recent years, the task and performance requirements of individual robots have rapidly increased in complexity and sophistication. Therefore, it has become difficult to meet these demands efficiently. Collaborative operations using MAS have gained attention as solutions to this problem [2].

H. Kimura (✉) · A. Okuyama
Tokai University, Kitakaname 4-1-1, Hiratsuka, Kanagawa, Japan
e-mail: 2CTAD001@mail.u-tokai.ac.jp

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_39

Previous MAS studies have assumed that the graph is always connected; that is, communication paths always exist between agents. However, if agents are assumed to be actual robots, the communication range must be considered finite, because information is exchanged among agents through mutual communication using wireless communication devices. In this case, the agents may fall into a disconnected condition, which is, there is no communication path between agents. Therefore, a control method for MAS that considers the disconnected condition is required.

In this case, there exists a case in which consensus cannot be achieved by conventional average consensus control for MAS considering disconnected conditions depending on the initial condition. To address this problem, we designed a controller including variable gain and demonstrated its validity in achieving consensus [3]. However, this variable gain was obtained through trial and error for the required conditions, and a systematic theory has not yet been established.

Therefore, we focused on funnel control to address this issue. Funnel control is a type of high-gain adaptive control, proposed in [4]. Its control objective is to converge control deviation within the predefined function called funnel function $\mathcal{F}_\varphi$ which is defined to be satisfied by the transient response.

In this study, we designed a consensus control method based on funnel control to achieve a consensus for MAS considering the disconnected conditions. In addition, simulations were performed to verify their effectiveness.

2 Consensus Problem

2.1 Algebraic Graph Theory

A graph comprises vertices and edges that connect them. Let N be the number of agents, where we denote $\mathcal{V} = \{1, 2, \ldots, N\}$, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, and $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ as the vertex set, edge set, and graph, respectively. We assume two arbitrary vertices i and j, the set of vertices that are adjacent to vertex i is called the neighborhood and is given by

$$\mathcal{N}_i \triangleq \{j \in \mathcal{V}, (i, j) \in \mathcal{E} \text{ and } j \neq i\} \tag{1}$$

For graph $\mathcal{G}$, the matrix expressing the adjacency is called adjacency matrix $\mathcal{A}$. The elements of $\mathcal{A} = \left[a_{ij}\right] \in \mathbb{R}^{N \times N}$ are given by the following expression:

$$a_{ij} \triangleq \begin{cases} 1, & (i, j) \in \mathcal{E} \text{ and } i \neq j \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

Adjacency matrix $\mathcal{A}$ is symmetric for an undirected graph. Moreover, the matrix with an in-degree as a diagonal element is called the degree matrix $\mathcal{D} = \mathrm{diag}\left\{d_1^{in}, d_2^{in}, \ldots, d_N^{in}\right\} \in \mathbb{R}^{N \times N}$, and its elements are given as follows:

$$d_i^{\text{in}} \triangleq \sum_{j=1}^{N} a_{ij} \tag{3}$$

The matrix based on adjacency matrix $\mathcal{A}$ and degree matrix $\mathcal{D}$ is called the graph Laplacian $\mathcal{L} \in \mathbb{R}^{N \times N}$ and is defined by the following equation:

$$\mathcal{L} \triangleq \mathcal{D} - \mathcal{A} \tag{4}$$

In MAS, the graph Laplacian $\mathcal{L}$ represents the overall system characteristics. The second-smallest eigenvalue (of $\mathcal{L}$) is a key factor in determining the consensus speed of the graph. Similar to $\mathcal{A}$, $\mathcal{L}$ is symmetric in an undirected graph.

2.2 Consensus Problem and Consensus Control

In this study, MAS was assumed to be a discrete-time system. The state variable of agent i for time k is denoted by $z_i[k] \in \mathbb{R}^{2\times 1}$ and is expressed as follows:

$$z_i[k] = \begin{bmatrix} x_i[k] & y_i[k] \end{bmatrix}^T (i = 1, \ldots, N) \tag{5}$$

where $x_i[k] \in \mathbb{R}$ and $y_i[k] \in \mathbb{R}$ are the coordinate points of x and y axes, respectively. The state variable collectively represents all agents, as $z[k] = \begin{bmatrix} z_1[k] & \cdots & z_N[k] \end{bmatrix}^T \in \mathbb{R}^{2N\times 1}$. The dynamics of agent i are given by the following differential equation:

$$z_i[k+1] = z_i[k] + T_S u_i[k], z_i[0] = z_{0i} \tag{6}$$

where $u_i[k] \in \mathbb{R}^{2\times 1}$ denotes the control input, $z_{0i} = \begin{bmatrix} x_{0i} & y_{0i} \end{bmatrix}^T$ denotes the initial condition, and T_S denotes the sampling time. The control input $u_i[k] \in \mathbb{R}^{2\times 1}$ is given by the following expression:

$$\begin{aligned} u_i[k] &= \varepsilon \sum_{j \in \mathcal{N}_\text{i}} \left(z_j[k] - z_i[k]\right) \\ &= \varepsilon \sum_{j=1}^{N} a_{ij} \left(z_j[k] - z_i[k]\right) \end{aligned} \tag{7}$$

where ε is the control gain.

Consensus refers to the asymptotic convergence of the state variable of all agents through information exchange within neighborhoods. A consensus is considered to have been achieved when the following equation holds for arbitrary agents i and j:

$$\lim_{k\to\infty} \left(z_j[k] - z_i[k]\right) = 0 \tag{8}$$

Furthermore, with respect to constant α, when the following equation is satisfied, α is called the consensus value:

$$\lim_{k \to \infty} z_i[k] = \alpha(\forall i = 1, 2, \ldots, N) \tag{9}$$

If the graph is connected, α is uniquely determined. In the case of the consensus problem with (7), α is obtained from the average value of the initial states of all the agents. This is called the average consensus problem, and the consensus value α is given by:

$$\alpha = \frac{1}{N} \sum_{i=1}^{N} z_{0i} \tag{10}$$

Thus, Eq. (7), which is the consensus control method in the case of MAS with a fixed graph, can always achieve consensus, and its consensus value can be obtained from the initial states.

3 Extension for MAS Considering Limited Communication Range

In this section, with respect to the basic MAS using (6) and (7), we describe the extended consensus control when the communication range of the agent is limited.

Figure 1 shows an omni-directional mobile robot that was assumed to be an agent in this study. This robot has an omni-wheel and can move in any direction without changing its own posture by adjusting the output of each wheel. For simplicity, the agents are represented as a mass model, and the state variable of (5) is the center of the robot as shown in Fig. 1. Considering the agents to be actual robots, they communicate with each other using wireless communication devices. Therefore, its communication range must be considered finite, and it is necessary to extend MAS by considering them.

In this study, the communication range was set as concentric circles of radius r centered on each agent. Thus, the elements of the time-invariant adjacency matrix in the basic MAS shown in (2) become time-varying. Based on this communication range, $\mathcal{R}_i$, (2), (3), and (4) can be rewritten as follows:

$$a_{ij}[k] \triangleq \begin{cases} 1, & z_j[k] \in \mathcal{R}_i[k] \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

$$d_i^{in}[k] \triangleq \sum_{j=1}^{N} a_{ij}[k] \tag{12}$$

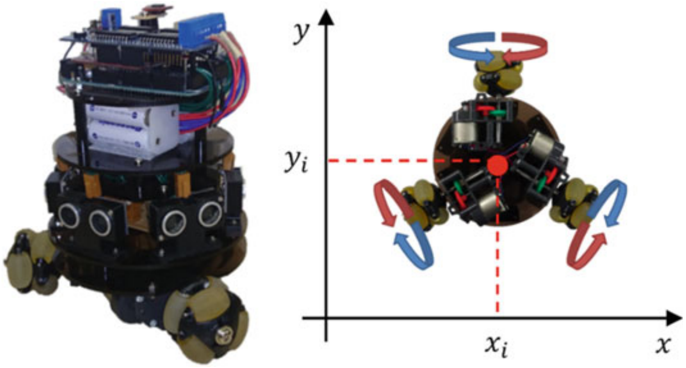


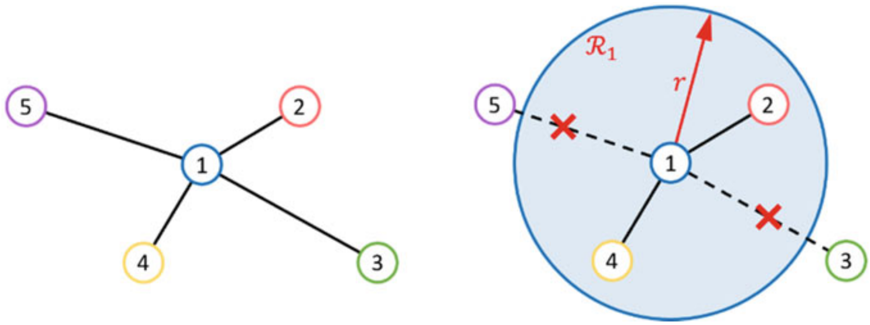
Fig. 1 Omni-directional mobile robot

$$\mathcal{L}[k] \triangleq \mathcal{D}[k] - \mathcal{A}[k] \tag{13}$$

In the above equation, the elements of the adjacency matrix are 1 if they lie within the communication range of agent i and 0.

Figure 2 shows examples of the graph considering communication range $\mathcal{R}_i$. Figure 2a, b shows the graphs based on (4) and (13), respectively.

In Fig. 3a, the adjacency matrix, degree matrix, and graph Laplacian are obtained as follows:



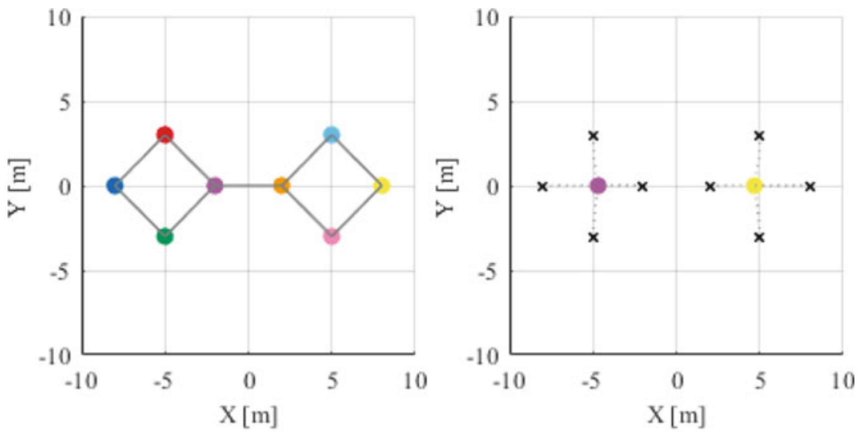
(a) Graph based on (4)

(b) Graph based on (13)

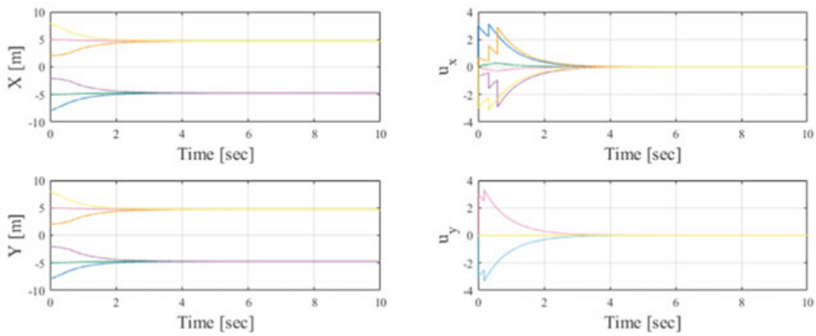
Fig. 2 Example of the graph considering communication range $\mathcal{R}_i$

$$\mathcal{A} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \mathcal{D} = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathcal{L} = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix}$$



(a) Initial condition (b) Result trajectory



(c) Time-history response of state variables (d) Time-history response of inputs

Fig. 3 Simulation result with conventional consensus control

Moreover, in Fig. 3b, the adjacency matrix, degree matrix, and graph Laplacian are obtained as follows:

$$\mathcal{A} = \begin{bmatrix} 0\,1\,0\,1\,0 \\ 1\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0 \\ 1\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0 \end{bmatrix}, \mathcal{D} = \begin{bmatrix} 2\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0 \\ 0\,0\,0\,1\,0 \\ 0\,0\,0\,0\,0 \end{bmatrix}, \mathcal{L} = \begin{bmatrix} 2 & -1 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

In Fig. 3b, Agents 3 and 5 are out of the communication range of Agent 1, and the elements of the adjacency matrix a_{13}, a_{15}, a_{31}a and a_{51}, which represent adjacency between them, go from 1 to 0. Therefore, the graph Laplacian $\mathcal{L}$ becomes a time-variant, and the overall system also becomes time-variant. Because the system becomes time-varying, the control input in (7) is extended as follows:

$$u_i[k] = \varepsilon_i[k] \sum_{j=1}^{N} a_{ij}[k]\big(z_j[k] - z_i[k]\big) \tag{14}$$

Here, $\varepsilon_i[k]$ is the control gain expressed as a vector for each agent. Henceforth, MAS with a limited communication range based on (14) is referred to as the conventional method.

Figure 3 shows an example of the simulation results obtained by consensus control using the conventional method. Figure 3a–d shows the initial state, resulting trajectory, time-history response of state variables, and time-history response of inputs by consensus control using the conventional method, respectively. In this simulation, we set $\varepsilon_i[k] = 1/d_i^{\text{in}}[k]$. Because the initial state in Fig. 3a is connected, a consensus is always achieved for the basic MAS in (7). However, it was confirmed that the agents were divided into two connected components, as shown in Fig. 3b, which resulted from applying the conventional method. Thus, there are cases in which consensus cannot be achieved using MAS with a limited communication range. Therefore, a control method is required to solve this problem.

4 Funnel Control

Funnel control is a high-gain adaptive control law that aims to converge the tracking deviation between the output of the control target and the reference value within a predefined performance funnel [4]. Figure 4 shows the concept of performance funnel $\mathcal{F}_\varphi$. $e(t)$ and $\varphi(t)$ are the deviations between the reference value and the current value and the scalar function, respectively. $\varphi(t)$ is called the funnel boundary and is set as a function that satisfies the following conditions.

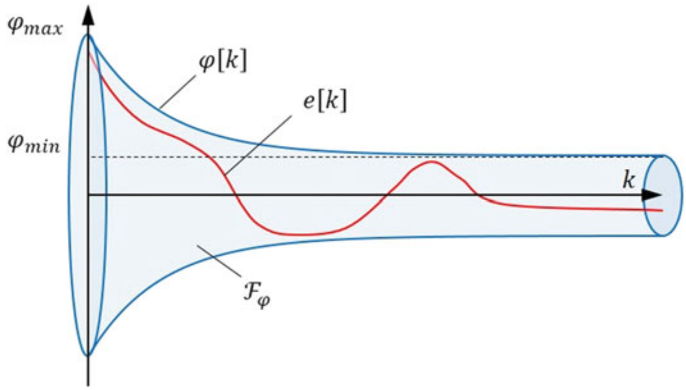


Fig. 4 Concept diagram of performance funnel $\mathcal{F}_\varphi$

$$\varphi : [0, \infty) \to [\varphi_{\min}, \varphi_{\max}](\varphi_{\max} > \varphi_{\min} > 0) \tag{15}$$

$\varphi_{\max}$ and $\varphi_{\min}$ are the maximum and minimum values of function $\varphi(t)$, respectively. In [5], the funnel boundary $\varphi(t)$ was chosen as follows:

$$\varphi(t) = \varphi_{\min} + (\varphi_{\max} - \varphi_{\min})\exp(-\lambda t) \tag{16}$$

where λ and t are the time constant and time in the continuous-time system, respectively. Thus, funnel boundary $\varphi(t)$ starts at $\varphi_{\max}$ and converges exponentially to $\varphi_{\min}$. $\mathcal{F}_\varphi$ is called a funnel function and set as a function that satisfies the following conditions:

$$\mathcal{F}_\varphi \triangleq \{(t, e) | \|e\| < \varphi(t)\} \subseteq [0, \infty) \times \mathbb{R} \tag{17}$$

The objective of funnel control is to converge the deviation $e(t)$ within the funnel function $\mathcal{F}_\varphi$ as shown in Fig. 4. Therefore, it was proposed to increase the controller gain γ_φ as the deviation approaches the funnel boundary $\varphi(t)$ [4]. The simplest example as follows:

$$\gamma_\varphi(t, e) = \frac{\delta}{\varphi(t) - \|e(t)\|} \tag{18}$$

Here, δ is an arbitrary constant.

5 Consensus Control Based on Funnel Control Considering Disconnected Conditions

This section describes a solution to the problem of the inability to achieve consensus in MAS with the limited communication range presented in Sect. 3. We propose a consensus control method with a newly designed control gain in (14) based on funnel control.

The control gain is defined as $\varepsilon^{\varphi}[k] = \left[\varepsilon_{ij}^{\varphi}[k]\right] \in \mathbb{R}^{2N \times 2N}$ to distinguish it from the control gain of the conventional method. As shown in (14), in conventional consensus control, the direction of motion and speed of the agents are determined by the number of agents in the neighborhood and the distance between them. Therefore, the required control gain performance must compensate for these effects. The distance between agents i and j is defined as follows:

$$e_{ij}[k] = z_j[k] - z_i[k] \tag{19}$$

The range of $e_{ij}[k]$ is $0 \le e_{ij}[k] \le r$ based on the radius r of the communication range. From the above, the funnel function $\varphi[k]$ is prepared for each relationship between agents i and j and is defined as follows:

$$\varphi_{ij}[k] = \varphi_{\min} + (\varphi_{\max} - \varphi_{\min})\exp\Big(-\big(\lambda\big(k \times T_S - \tau_{ij}[k]\big)\big)^2\Big) \tag{20}$$

$$\tau_{ij}[k] = \begin{cases} k \times T_S, & a_{ij}[k] > a_{ij}[k-1] \\ \tau_{ij}[k-1], & \text{otherwise} \end{cases} \tag{21}$$

Here, $\tau_{ij}[k]$ is an arbitrary time updated only when agents i and j are newly connected. Figure 5 shows the proposed performance funnel using (20).

From the above, the control gain $\varepsilon^{\varphi}[k]$ is defined as follows:

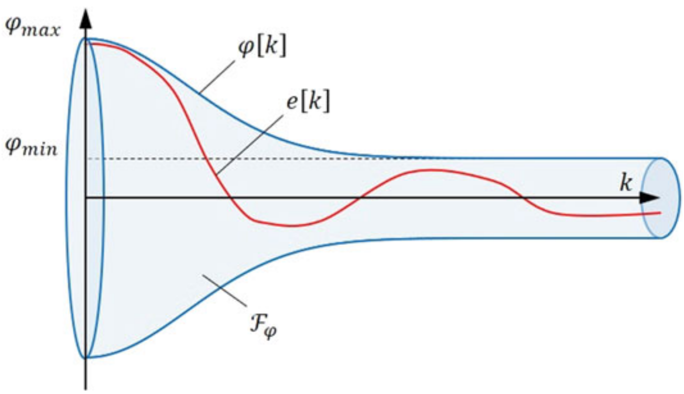


Fig. 5 Concept diagram of proposed performance funnel $\mathcal{F}_{\varphi}$

$$\varepsilon_{ij}^{\varphi}[k] \triangleq \begin{cases} \frac{\delta}{\left(d_i^{in}[k]\right)^2} \frac{1}{\varphi_{ij}[k]-e_{ij}[k]}, & a_{ij}[k]=1 \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

$$\varepsilon_{ii}^{\varphi}[k] \triangleq \frac{1}{d_i^{in}[k]} \sum_{j=1}^{N} \varepsilon_{ij}^{\varphi}[k] \tag{23}$$

Based on the above, we propose the following control method expression using the newly proposed control gain $\varepsilon^{\varphi}[k]$:

$$u_i[k] = \sum_{j=1}^{N} a_{ij}[k]\varepsilon_{ij}^{\varphi}[k]\big(z_j[k] - z_i[k]\big) \tag{24}$$

The input collectively represents all agents, as $u[k] = \left[u_1[k] \cdots u_N[k]\right]^T \in \mathbb{R}^{2N\times 1}$ and is obtained as follows:

$$u[k] = -\varepsilon^{\varphi}[k] \odot \mathcal{L}[\mathrm{k}]\mathrm{z}[\mathrm{k}] \tag{25}$$

Here, $\odot$ represents Hadamard operation and is used as follows:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix} \odot \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1N} \\ b_{21} & b_{22} & \cdots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \cdots & b_{NN} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1N}b_{1N} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2N}b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}b_{N1} & a_{N2}b_{N2} & \cdots & a_{NN}b_{NN} \end{bmatrix}$$

Henceforth, the MAS with a limited communication range based on (24) is referred to as the proposed method.

6 Simulation Study

In this section, considering MAS with a limited communication range, a simulation is performed to verify the effectiveness of consensus control based on the method proposed in Sect. 5.

6.1 Simulation Overview

Figure 6 shows an example of the initial conditions used in the simulation. Each circle represents an agent, and the gray lines represent the communication paths between agents. In the preliminary stage, a simulation using the conventional method was conducted for the randomly prepared initial conditions. The initial conditions under which consensus could not be achieved were adopted as the initial conditions, as shown in Fig. 5b. Ten thousand such initial conditions were prepared. Thus, not all initial conditions can achieve a consensus with conventional methods that do not consider a limited communication range. Table 1 lists the parameters used in these simulations.

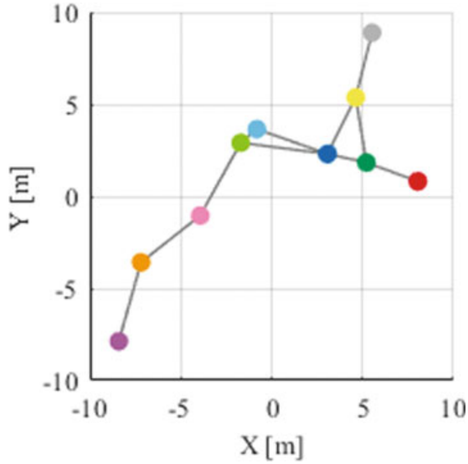


Fig. 6 Example of the initial conditions

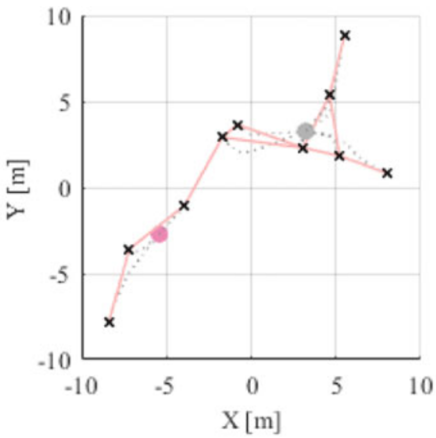
Table 1 Parameters for simulations

Symbol	Meaning	Value
N	Number of agents	10
F	Field size	20×20 m^2
r	Radius of communication range	5 m
T_{end}	Simulation time	30 s
T_S	Sampling time	1 ms
$\varphi_{\max}$	Maximum value of $\varphi[k]$	$1.1 \times r$
$\varphi_{\min}$	Minimum value of $\varphi[k]$	$0.3 \times r$
λ	Time constant of $\varphi[k]$	0.5 s
δ	Constant value	5

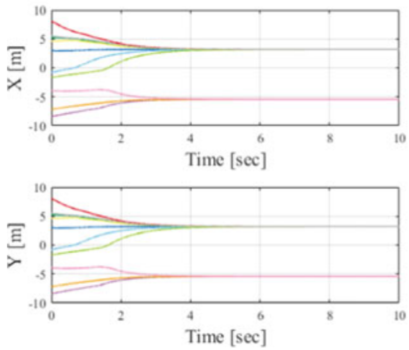
6.2 Simulation Results

Figures 7 and 8 show the simulation results for the conventional and proposed methods, respectively. Figures 9a and 10a show the movement trajectories of the agents. The circles, black x marks, red lines, and grey dotted lines indicate the agents, initial values, communication paths in the initial state, and movement trajectories, respectively. These results confirm that the proposed method can achieve consensus even under the initial conditions where consensus cannot be achieved using the conventional method.

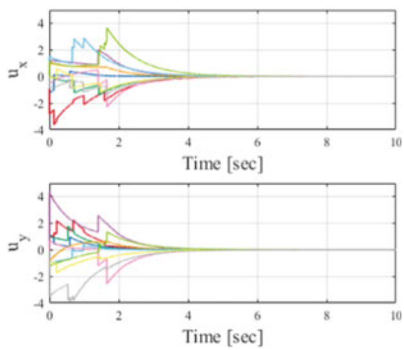
Figures 9b and 10b show the time-history responses of the state variables. Figure 9b, which shows the result of the conventional method, shows that the convergence value of the state variable splits into two at less than 5 s. In contrast, Fig. 10b,



(a) Result trajectory

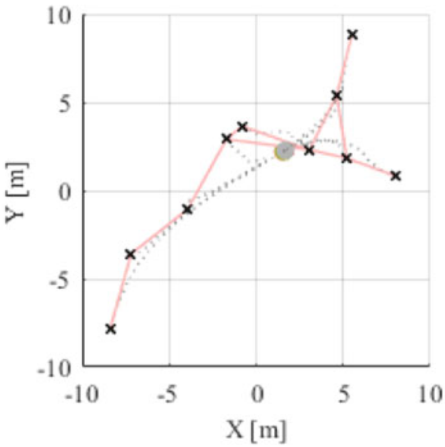


(b) Time-history response of the state vari ables

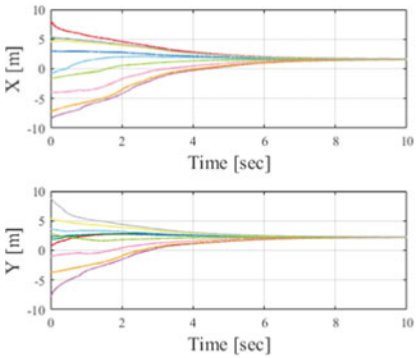


(c) Time-history response of inputs

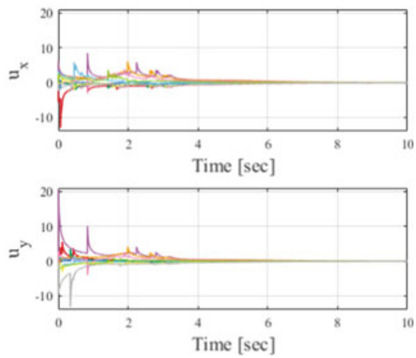
Fig. 7 Simulation results using the conventional control method which enlarge the part up to 10 s



(a) Result trajectory



(b) Time-history response of the state variables



(c) Time-history response of inputs

Fig. 8 Simulation results using the proposal control method

which shows the result of the proposed method, shows that the convergence value of the state variables does not split, and all agents converge at approximately 7 s.

Figures 9c and 10c show the time-history response of the inputs. Figure 9c shows the result of the conventional method and the input range of ±4. Figure 10b, which is the result of the proposed method, shows that the inputs ranged from approximately −15 to 20, and the range decreases after 1 s.

Figure 11 shows the time-history response of the deviations between Agent 1 and each agent and the performance funnel determined in (20). As shown in Fig. 11, each performance funnel was generated when Agent 1 and another agent were newly connected. Moreover, all the deviations were suppressed inside each funnel function.

Simulations were performed for 10,000 different initial conditions, and the results confirmed that a consensus was achieved in all cases. Figure 11 shows the results obtained using the proposed method with other initial conditions, and it is confirmed

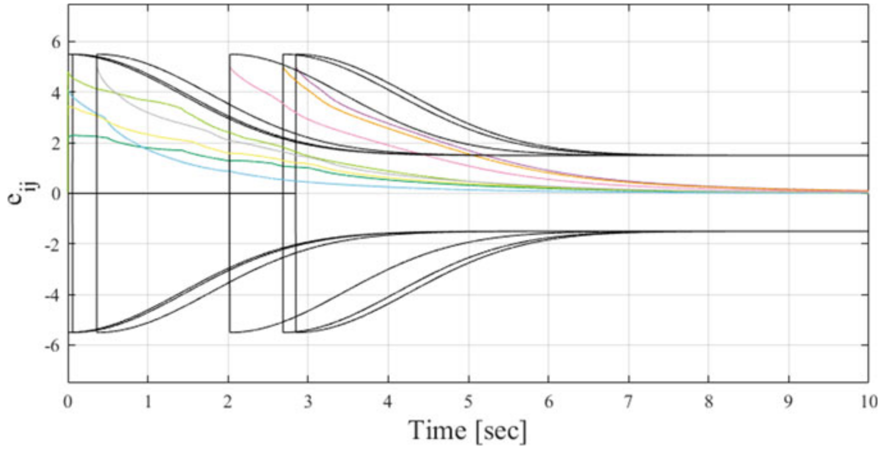


Fig. 9 Time-history response of deviation for Agent 1 and the performance funnel which enlarge the part up to 10 s

that a consensus is achieved even if the distribution of agents is biased. Therefore, the proposed method with the conditions listed in Table 1 was effective for MAS, considering the limited communication range.

However, the proposed method has parameters, namely: $\varphi_{\max}$, $\varphi_{\min}$, λ, and δ that need to be adjusted for each control object. $\varphi_{\max}$ and $\varphi_{\min}$ are the upper and lower limits of funnel function $\varphi(t)$, respectively. $\varphi_{\max}$ affects the convergence in situations where the distance between agents is large, and $\varphi_{\max}$ affects the convergence speed where the distance between agents is small. λ determines the degree of convergence of $\varphi(t)$, and δ determines the volume of the control gain $\varepsilon_{ij}^{\varphi}[k]$. Therefore, the graph may split, or the convergence speed may slow do if the value of the parameters is not appropriate. From the above, establishing a theoretical proof for the value of the provides scope for further research.

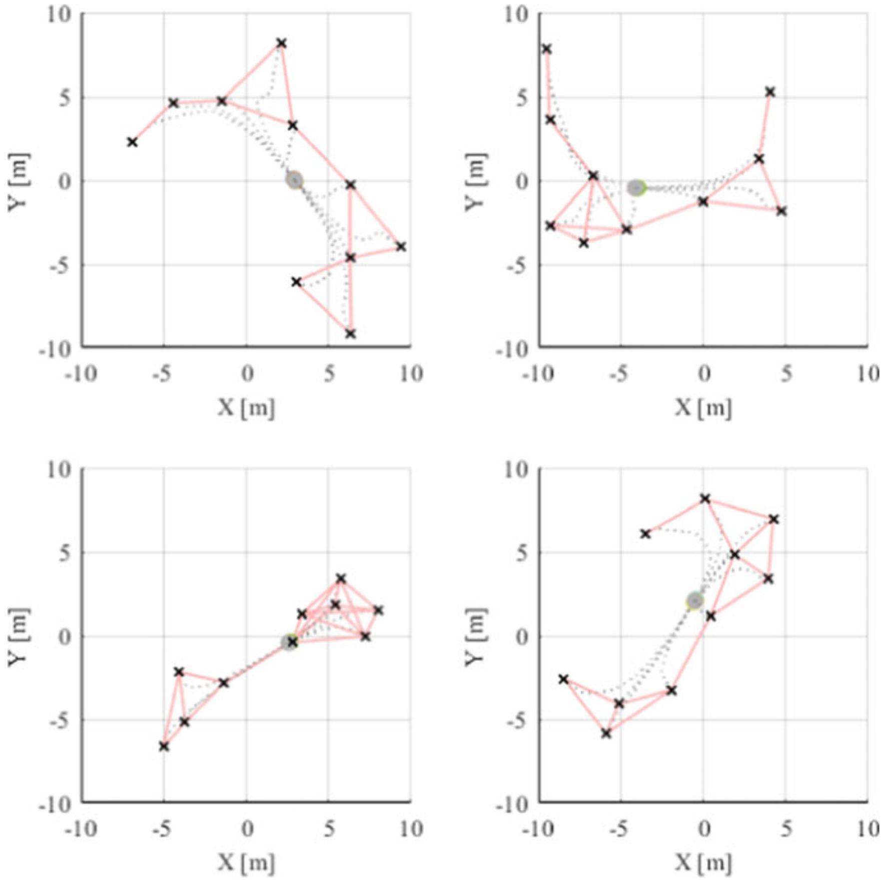


Fig. 11 Simulation results using the proposed consensus control with other initial conditions which enlarge the part up to 10 s

7 Conclusion

In this study, we designed a consensus control method based on funnel control that achieved consensus for MAS considering the disconnected conditions and performed simulations to verify the effectiveness of our method. The results indicate that a consensus was achieved in all different initial conditions, which emphasizes the effectiveness of the proposed method for MAS with a limited communication range. Thus, it was shown that the MAS was extended to take into account the real environment, which depend on the communication range of each agent.

Simulations were performed for 10,000 different initial conditions, and the results confirmed that a consensus was achieved in all cases. Therefore, it was confirmed that the proposed method, funnel function, and control gain designed in (20)–(23) were not affected by the bias of the distribution of initial conditions. However, the parameters must be readjusted when the communication range is changed. Thus, the establishment of an adjustment method with theoretical proof is an issue.

In future, we will clarify the performance limit of the parameters and establish a theoretical proof for the adjustment method of the values. In addition, the simulations in this study were carried out assuming a real robot; however, the model remains a mass model. Therefore, the simulations in omni-directional mobile robot shown in Chap. 3 are carried out to extend the MAS to a more realistic environment.

References

1. Azuma S, Nagahara M, Ishii H, Hayashi N, Sakurama K, Hatanaka T (2015) Control of multi-agent systems. Corona Publishing co. ltd., Japan, vol 1
2. Moshtagh N, Jadbabaie A (2007) Distributed geodesic control laws of flocking of nonholonomic agents. IEEE Trans Autom Control 52(4):681–686
3. Kimura H, Okuyama A (2017) Average consensus control in a multi-agent system with communication restriction. In: Micromechatronics. Japan
4. Ilchmann A, Ryan EP, Sangwin CJ (2002) Tracking with prescribed transient behaviour. ESAIM Control Optim Calculus Variat 7:471–493
5. Shim H, Trenn S (2015) A preliminary result on synchronization of heterogeneous agents via funnel control. In: 54th IEEE conference on decision and control, pp 2229–2234

A Secure and Effective Solution for Electronic Health Records with Hyperledger Fabric Blockchain

Doruntina Nuredini, Daniela Mechkaroska, and Ervin Domazet

Abstract Blockchain is a decentralized database where you are allowed to add blocks and ready chains anytime. Blockchain technology has a great potential varying from the financial sector, including the IoT, e-commerce, accounting and auditing, electronic voting, asset and identity management, supply chain management, taxation, telecommunications, health care, and public services. Security and efficiency of these types of applications remain an issue. Blockchain can address all of these concerns in an efficient way, where it can provide benefits including easy deployment and maintenance, seamless authentication, privacy, and security. Blockchain serves as the basis for many applications in smart cities, whereas in this paper, we will put the focus, particularly on the sub-area of smart cities, namely the health-care sector. We will make use of the experiences we had in smart cities, in order to suggest an efficient solution for electronic health records (EHRs). In this paper, we suggest a framework for EHRs that makes use of blockchain technology to store medical records more effectively, securely, and reliably while also facilitating quick access to them. In this regard, we are considering using Hyperledger, a permissioned blockchain platform.

Keywords Blockchain · Electronic health records · Security · Efficiency · Hyperledger

D. Nuredini (✉) · D. Mechkaroska
University of Information Science and Technology "St. Paul the Apostle", Ohrid, Macedonia
e-mail: doruntina.nuredini@uist.edu.mk

D. Mechkaroska
e-mail: daniela.meckaroska@uist.edu.mk

E. Domazet
International Balkan University, Skopje, Macedonia
e-mail: ervin.domazet@ibu.edu.mk

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_40

1 Introduction

Smart cities are being developed thanks to the advances in the field of Internet of things (IoT)-based applications. Transportation, business, industry, health care, smart homes, and finance are just a few sophisticated services that smart cities provide. While enhancing the quality of life for citizens, these applications require extremely high levels of security for data management. We can utilize blockchain to develop smart cities with improved security and privacy.

Blockchain is a decentralized database, having the option to read the chains and related blocks anytime, from anywhere, where only the addition of blocks to the chains is allowed, and all the other operations are strictly permitted. The Internet of things (IoT) e-commerce, accounting and auditing, electronic voting, asset, and identity management, supply chain management, taxation, telecommunications, health care, and public services are just a few of the industries that could benefit from using blockchain technology.

Smart cities are expected to be improved in the sense of becoming more interconnected, instrumented, intelligent, livable, safe, sustainable, and resilient as a result of recent advancements in niche areas of ICT, including ML, AI, blockchain, automation, etc. Security, privacy, and transparency are always the most critical aspects of any software. These all can be addressed by blockchain. Smart city transactions can be recorded in a blockchain in various formats. Smart contracts allow for the automatic interchange of data and the execution of complex legal processes. Due to smart contracts and decentralized applications, blockchain offers a high degree of autonomy to perform smart transactions during the operation of a smart city. The benefits of blockchain technology include simple implementation and maintenance, seamless authentication, privacy, and security.

Blockchain serves as the basis for many applications in smart cities, whereas in this paper, we will focus on the healthcare sector. We will try to elaborate on our experience with smart cities and discuss how that can be transferred to EHRs. We are going to discuss the challenges and benefits of the adoption of blockchain within health care, as well as technical requirements for block chain-based healthcare applications will be mentioned. In this paper, we suggest a framework for electronic health records (EHRs) that makes use of blockchain technology to store medical records more effectively, securely, and reliably while also facilitating quick access to them. We used Hyperledger [1], a permissioned blockchain platform.

2 Internet of Things (IoT)

Today we live in the age of intelligent technologies, representing pervasive computing or Web 3.0 [2]. The IoT has become a richer field to express this new technology. The phrase is a combination of the special words "Internet" and "thing", where the former is "the worldwide network of interconnected computer networks based on a standard communication protocol", and the latter is a "virtual network, real, moving or stationary object that constantly transmits information to other objects". The phrase "Internet of things" refers to the extension of the Internet connection to the most comprehensive range of object kinds. Remote monitoring and management of items are possible thanks to the Internet's ability to interchange and communicate the data that specialized sensors collect [3].

2.1 IoT Applications for Smart Cities

According to the study [4], 50 billion products will be connected to the Internet by 2020. Furthermore, statistics show that (approximately) six products are used per person [5]. This means that the use of the IoT will be enormous, and the performance will be six times more efficient. This number itself is sufficient to demonstrate the future expansion of IoT technology.

The building block of smart cities is IoT technology. There are many IoT options for smart cities, ranging from Internet-connected garbage cans and connected buildings to IoT-based fleet management and Internet-connected cars. City authorities can remotely monitor and control IoT-related equipment for smart cities to ensure efficient operation.

Other examples of crucial IoT applications for a smart city include smart lighting systems, smart traffic management, IoT-based smart waste management in cities, IoT-based transportation, health care, etc.

A few key benefits of IoT applications for smart cities include increased efficiency and effectiveness, reduced crime, better services, reduced traffic congestion, and collection of large amounts of data on various aspects of city functioning. In addition, IoT performs urban infrastructure and environmental management activities. When developing applications for smart cities, sensors are usually used in large numbers and need to be connected. This raises specific challenges related to the construction of efficient and secure communication infrastructures, including location information and data collection [6–8].

Within the smart city, digital services are more advanced than traditional legacy-based networks. This, in turn, requires the most recent technologies, such as considering digital and communication technologies of the information to improve operations and overall services to its residents [9]. In this paper, we focus on IoT applications for health care, particularly smart cities.

Even in the age of smart cities, the most frequently compromised patient data includes names, contact information, and descriptions of diseases. These specifics are digitally preserved on a system known as an electronic health record (EHR). All information is stored in the EHR, including information on the patient, scan reports, clinical notes, sensor data, billing data, medications, medical history, insurance information, and other relevant details. Future medical research that aims to enhance patient care and clinical practice efficiency might benefit from the EHR.

This data is not accessible to patients and their caregivers, but it is easily accessible to unauthorized third parties and easily attacked by hackers. This creates an imbalance in data *accessibility, privacy,* and *security*.

Instead of the relevant patients, healthcare organizations have been in charge of managing the EHR. This makes it difficult for other medical facilities to access patient information to give patients the best possible medical advice. As a result, patients must save their health information for access in the future.

The blockchain makes the transaction decentralized and provides an immutable ledger. Three essential characteristics of the blockchain are security, transparency, and decentralization. These crucial components ensure that the system is highly secure, prevent data modification, and restrict access to only authorized users.

2.2 Blockchain Technology in Smart City

With its features, blockchain technology offers cutting-edge answers to the primary issues that smart cities face [10]. Due to this feature, blockchain allows for fully digitalized cities where everything is regulated digitally, which lowers human labor and saves much time. In order to foster trust, enhance infrastructural convenience, and increase operational efficiency in a smart city, blockchain, in conjunction with IoT, helps to promote transparency, privacy, security, and data integrity. According to a study [11], the blockchain can be used in several smart city industries, including health care, transportation, education, energy use, waste management, agriculture, etc.

2.3 Healthcare Blockchain Applications in Smart Cities

The basis of a happy life for a citizen is health. The development of medical technology has brought significant benefits to the public [12]. Traditional health care cannot meet the needs of the exponentially growing world population. It is essential to transform legacy health care into smart, efficient, and sustainable architectures, due to the conflict between growing demand and limited resources. Wearable technology, smart hospitals, emergency services, and ambulance systems are some elements involved in making smart care a reality [13].

To enable patients to receive adequate care, patient data is essential. In the field of smart health care, efficient exchange of patient data between can help caregivers continuously check patients' condition and make real-time decisions regarding patient health, even if they are far from the patient. There are several advantages to using blockchain in smart assistance [14, 15]. For example, the blockchain can be used:

- To record medical data securely and irrevocably

- Patients have flexible access control and control over how their medical information is used
- Transparent supply chain
- Commitments applied through smart contracts.

3 Electronic Health Records with Hyperledger Fabrik Blockchain

In order to safely and securely store and ensure ease of access of the patient records, here we will propose a system for storing EHRs that uses blockchain technology.

A distributed ledger based on the blockchain was created with the support of the open-source initiative known as Hyperledger. One of the most well-liked Hyperledger projects is Hyperledger Fabric [1]. It is a blockchain infrastructure with permissions for creating tools, programs, and apps.

3.1 How Does Hyperledger Fabric Work?

In the Hyperledger Fabric network, there are different organizations that communicate with each other on the network. Each organization has a fabric *certificate authority* and *one or more peers*. Each organization uses an *ordering service* shared by the Fabric network, which helps process transactions on the network [1].

An organization-unique root certificate serves as a definition of that organization within a network. This certificate is generated based on the root one. It ensures that other entities on the network can connect users to their organizations. Additionally, entities (peer nodes) within that organization can easily be identified. Permissions for each entity on the network are also specified in these certificates, read-only permission or full channel access permission.

The certification authority maintains a root certificate for an organization. The certification authority also handles other related tasks and issues certificates to users within an organization.

To carry out duties on its behalf, an organization constructs one or more peer nodes as components. Each peer node keeps a local copy of the ledger for access, stores and executes the smart contract code (a chaincode in Fabric), and approves the proposed transactions from the network. To read the ledger, introduce a new chain code into the network, or propose a new transaction, and fabric clients often communicate with peer nodes [1].

The ordering service ensures that newly added transactions to the network are appropriately approved and sorted into new blocks. A new transaction block is then sent by the ordering service to peer nodes within each organization. With this new block, peer nodes update their copies of the local ledger [1].

3.2 High-Level Architecture of the Proposed EHR Framework

The architecture of our proposed framework is based on the architecture of the Hyperladger Fabric model, as we already mentioned. Figure 1 presents the high-level architecture of the EHR framework. The application is aimed at three different roles: *patients*, *doctors,* and *administration*. The administration role aims to serve as a controller that can register other roles.

In this framework, only the patient and the doctor as authorized parties can access the data from the blockchain. The latter can do this if the former gives explicit permission.

Since this framework uses permissioned blockchain Hyperledger, doctors and patients should be registered to the *membership service* (MSP). After the registration process, they will have access to the system, and a certificate of authority will be created. An MSP specifically abstracts away the cryptographic protocols and methods involved in certificate issuance, certificate validation, and user authentication. An MSP is free to establish its definition of identity, as well as the guidelines for identity validation and authentication (signature generation and verification).

The certificate of authority is created. An MSP specifically abstracts away the cryptographic protocols and methods. Once the doctor and patient are registered, they are granted with specific *user id* and *password*, in order to have access to the system.

Each transaction in the network is encrypted with a public key that is generated by the MSP, and then participants in the network who are granted access to the transaction obtain the private key to be able to decrypt the transaction. Basically, these participants are provided with different types of privileges that allow them to do certain actions.

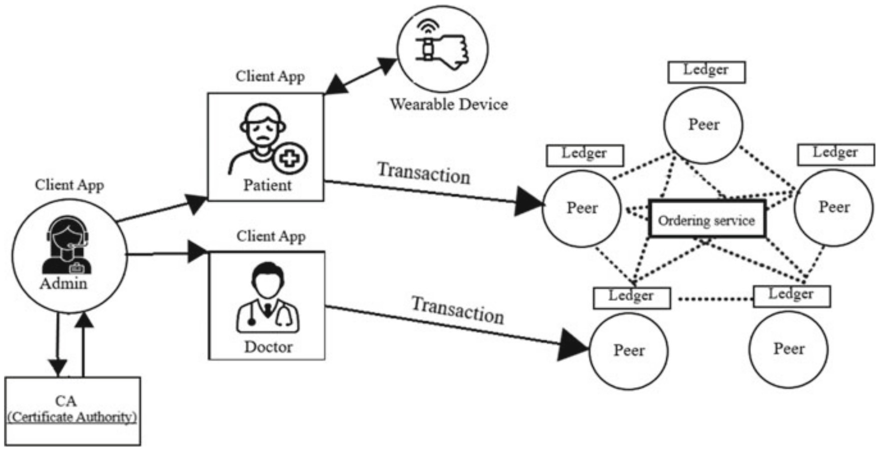


Fig. 1 High-level architecture of the proposed framework

The transaction is transmitted to all nodes and authenticated by the Hyperledger consensus algorithm.

4 Analysis and Discussion

Here are two reasons why Hyperledger (permissioned blockchain) is more superior over Ethereum (permissionless blockchain) for our framework:

It can be troublesome because anyone can join the network anonymously and without authorization in a permissionless blockchain. The identities of each network participant must be known in the case of an electronic medical record system. Utilizing Hyperledger, a permissioned blockchain network where each participant's identity is known makes sense as a result.

Medical data about patients is susceptible. Because every member of the network participates in a consensus mechanism in Ethereum, data stored on the distributed ledger that requires a higher level of privacy becomes available to all network users. Since only nodes allowed by authorities can view patient medical data thanks to Hyperledger's permissioned blockchain, the privacy needs of those data are correctly met [16].

Privacy is the biggest problem faced in permissionless blockchain platforms like Ethereum. When it comes to the maintenance and operation of the health system, the need to protect data from unauthorized persons is logical, which enables the privacy of the entire health system. Table 1 gives comparison of the features associated with the already existing Ethereum frameworks with the proposed framework.

Table 1 Comparison of the features of the Ethereum-based framework and the proposed framework

Feature	Ethereum	Proposed
Blockchain-based	✓	✓
Authentication	✓	✓
Identity management	✓	✓
Decentralized access	✓	✓
Integrity	✓	✓
Availability	✓	✓
Flexibility	✓	✓
Private (permissioned)	×	✓
Public (permissionless)	✓	×

5 Conclusion

Technological advances have opened many new opportunities. Many architectures are eligible to increase their security and efficiency by using the proper blockchain technology.

Blockchain's usage in smart cities is increasing day by day. Our focus in this paper was the healthcare applications within smart cities. In this regards, many new frameworks are proposed.

In this paper, we firstly analyzed the current state, and we stated the concerns of the current architectures. Finally, we proposed a framework that is based on Hyperledger Fabric blockchain for the EHRs in order to increase their security and efficiency.

In future work, we plan to implement a pilot EHR system that is based on the Hyperledger Fabric framework. We believe that this will make a significant impact on small countries like North Macedonia and will be a good case for other countries to start its adoption.

References

1. AWS (2022) What is hyperledger fabric? Available: https://aws.amazon.com/blockchain/what-is-hyperledger-fabric/
2. Techtarget, Web 3.0 (Web3) (2022) Available: https://www.techtarget.com/whatis/definition/Web-30
3. IGI Global (2022) What is Internet of Things (IoT). Available: https://www.igi-global.com/dictionary/internet-of-things-iot/43226
4. Malik A, Magar AT, Verma H, Singh M, Sagar P (2019) Detailed study of an internet of things (Iot). Int J Sci Technol Res 8(12)
5. Madakam S, Ramaswamy R (2014) Smart homes (conceptual views). IEEE Xplore. https://doi.org/10.1109/ISCBI.2014.21
6. ISO/IEC (2014–2015) Smart cities, preliminary report. Available: https://www.iso.org/files/live/sites/isoorg/files/developing_standards/docs/en/smart_cities_report-jtc1.pdf
7. Kumar TMV (2017) Smart economy in smart cities: international collaborative research: Ottawa, St. Louis, Stuttgart, Bologna, Cape Town, Nairobi, Dakar, Lagos, New Delhi, Varanasi, Vijayawada, Kozhikode, Hong Kong. Springer, Berlin
8. Shahrour I, Xie X (2021) Role of Internet of Things (IoT) and crowdsourcing in smart city projects. Smart Cities 4(4):1276–1292
9. Gade D (2021) Introduction to smart cities and selected literature review. Int J Adv Innov Res 6(2, Part 4):7–15
10. Rawat S, Sah A (2012) An approach to enhance the software and services of health care centre. IISTE 3:126–137
11. Sah A, Dumka A, Rawat S (2018) Web technology systems integration using SOA and web services. In: Handbook of research on contemporary perspectives on web-based systems; IGI Global: Hershey, PA, USA, pp 24–45
12. Collins FS (2015) Exceptional opportunities in medical science: a view from the national institutes of health. JAMA 313(2):131–132
13. Xie J, Tang H, Huang T, Yu FR, Xie R, Liu J, Liu Y (2019) A survey of blockchain technology applied to smart cities: research issues and challenges. IEEE Commun Surv Tutorial 21(3):2794–2830

14. Kuo TT, Kim HE, Ohno-Machado L (2017) Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 24(6):1211–1220
15. Mettler M (2016) Blockchain technology in healthcare: the revolution starts here. In: Proceedings of IEEE HealthCom'16, Munich, Germany, pp 1–3
16. Usman M, Qamar U (2020) Secure electronic medical records storage and sharing using blockchain technology. Procedia Comput Sci 174:321–327

Machine Learning Algorithms for Geriatric Fall Detection with Multiple Datasets

Purab Nandi and K. R. Anupama

Abstract Privacy of data is a significant concern, especially in the healthcare domain; it is paramount to abide by the existing privacy regulations. However, we require large datasets for training the ML/DL models. Fusion of similar or related data from multiple datasets taken from various sources worldwide will therefore be beneficial while retaining the anonymity of the test subjects. The impact of using multiple and different datasets for testing and training is yet to be studied. In contrast, individual datasets have been split randomly for testing and training in multiple research works; using separate datasets and studying the accuracies of ML algorithms is yet to be done. In this paper, we use multiple datasets for testing and different datasets for training, and we have studied the impact of using different datasets. We found that there was a massive drop in accuracy when varied datasets were used for testing and training, especially when the data collection methodology and demographics were different; in real-life scenarios, especially in relation to geriatric fall detection systems, the training data would be collected from much younger volunteers due to the risk factor involved when asking the elderly to fall. So the dataset and the demographics for training will be completely different from the end users. Hence this analysis of using multiple datasets gains immense importance.

Keywords Multiple datasets · Machine learning · Deep learning · Federated learning

P. Nandi (✉) · K. R. Anupama
Department of Electrical and Electronics Engineering, BITS Pilani, K.K Birla Goa Campus, Zuarinagar, Goa 403726, India
e-mail: p20200056@goa.bits-pilani.ac.in

K. R. Anupama
e-mail: anupkr@goa.bits-pilani.ac.in

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_41

1 Introduction

With advances in medical research, life expectancy has increased, while there has been an increase in nuclear families at the same time. Hence the percentage of the elderly who live on their own has risen over the last decade. Falling is one of the most damaging events an elderly person may experience. According to various studies conducted by WHO [1], falls represent the second major cause of accidental deaths worldwide, particularly among people 65 or older. The morbidity is very high; especially in the case of geriatrics over 80, residing in care homes. The percentage of elderly in care homes that experience at least one fall in a year is about 50% [1]. At the same time, 40% of them experience recurrent falls. The response time in getting aid is crucial and has serious consequences when delayed. The delayed response may lead to co-morbidities and permanent disabilities. There has been an enormous increase in research, especially among the elderly for over the last two years in the area of fall detection. During the past decade, a lot of work, primarily generated from the United States, China and Germany [2], has been concentrating on improving the performance of fall detection systems.

The extent to which the fall affects the elderly depends upon multiple factors, such as the activity and the posture of the person during and before the fall. While there are similarities in various kinds of falls, certain parameters will also vary with the type of fall. Other than the direction that the fall takes, another critical factor is the duration of the fall. Any episode such as sudden chest spasm or fainting causes a fall that may last for an extended duration. Injuries are also possible because of the surface as well as any obstacles; this might cause the elderly to experience an abrupt and hard fall. Some research shows that the subject's age and gender also play a role in fall kinematics.

While several papers are based on various public datasets collected over the years, how the data was collected, what sensors were used, and machine learning algorithms [3] that were used to predict using the sensed data are not available.

From the previous literature survey [2], it is obvious that the use of multiple benchmark datasets has not been considered; most literature only compares machine learning algorithms' performance by varying the features extracted from the same dataset. In most cases, these values are heuristically selected from test sample of particular dataset. Since a single dataset is used, the accuracy of the model in detecting falls in environments that are different from the training dataset is questionable. All work we have surveyed so far uses the same dataset for testing and training. So, it can be concluded that since the data for test and training is selected randomly, the data collected from the same person is used for testing and training.

In real-life scenarios, the model may be trained using data collected from a set of volunteers or publicly available datasets. The person(s) using the trained model will not be one of the volunteers. The age, gender and biological parameters of the person using the model may differ entirely from the volunteers. The question is, under such a condition will the model predict correctly? This question is what this paper tries to address.

The paper is organized as follows; Sect. 2 gives the background of work, Sect. 3 talks about the use of multiple datasets for testing and training ML algorithms, Sect. 4 gives the data collection methodology used by us as well as describes the public datasets used in this work. Section 5 presents the results and analysis, and we conclude in Sect. 6.

2 Background

Initial research in geriatric fall detection was focussed on wearable devices that could be placed indoors in a retirement home. Wearable devices are generally equipped with IMU sensors; IMU sensors use position and motion-based data to detect the orientation of the person, in case a fall occurs the sensor will be able to detect a negative acceleration, typically these devices are worn on the torso. This might be difficult for senior citizens as they may not be comfortable moving around with the sensor strapped to their chest. Recently, commercially available smartphones had in-built IMU sensors along with multiple other sensors that can be used to track the activities of that person. Data collection from these devices is more efficient due to the high computing power available in these devices. A survey paper that highlights the use of smartphones in fall detection is presented in [4]; smartphones generally are equipped with accelerometers, gyroscopes and magnetometers that can be used to sense falls. Among these sensors, the data collected from the accelerometer is the most significant, as it can be used to detect variation in acceleration and its integral can be used to detect movement and position. There are two types of algorithms currently available in the literature that are used for fall detection: a. threshold-based algorithms b. machine learning-based algorithms.

A torso-mounted bi-axial gyroscope is used for detecting falls using a threshold-based algorithm described in [5]. A total of 10 male volunteers with no health issues were used for simulating falls and the gyroscope signals were analyzed for each fall. Each volunteer was asked to perform eight types of falls each repeated three times. While ADL activities were recorded using eight elderly people, while geriatric people can be used for simulating ADL, the health risk involved in asking them to simulate falls is very high, hence most datasets including the one described in [5] use young people to simulate falls while the elderly are used for performing ADLs. Many of the ADL activities that were simulated are quite similar to falls, such as sitting down, standing up, lying down and getting up, and entering and exiting a car, are some of the ADL activities that give sensor values similar to falls. Bourke and Lyons [5] demonstrates that by using three thresholds related to angular velocity and acceleration in combination with the angle of the torso, a 100% specificity can be obtained; however as stated earlier, sensors that are worn on the torso affect the mobility of the elderly. This research also detects falls that have an acceleration that is higher than 6G, while hard falls may result in big changes in acceleration values whereas soft-falls will produce a minor change in acceleration ranging from 3 to 3.5G. If we use wrist-based wearable devices the change in acceleration maybe even lesser than

3G; therefore ML techniques are more robust when compared with threshold-based techniques. Multiple papers that are based on the four traditional ML algorithms; Naïve Bayes (NB), Kth nearest neighbour (KNN), support vector machine (SVM) and decision trees are available. These algorithms use data collected from waist-worn, torso-worn, wrist-worn and thigh-worn devices. Ramachandran and Karuppiah [3] provides a complete analysis of these algorithms. Recurrent neural networks were recently used to detect falls. In [6] an RNN architecture that uses accelerometer signals fed into two long short-term (LSTM) layers is described by the authors. The output of these layers was passed through two feed-forward neural networks. The second of these networks generated the probability of a fall has occurred. The model was trained and later evaluated using the URFD dataset [7], which contains accelerometer data which is taken from a sensor that was placed on the pelvis. The RNN algorithm produced an accuracy of 91.71%. The accuracy was improved to 98.57% by random rotation of the acceleration signal and retraining the model. The authors in [8] also use an RNN algorithm to detect falls using only the accelerometer data. The core of their neural network is a fully connected layer that processes the raw data, and this is followed by two LSTM layers and the final layer is also a fully connected layer. The authors have also used some normalization and dropout layers in their RNN model. The model was trained and evaluated using the SiSFall dataset [9]. The SiSFall dataset has accelerometer values sampled at 200 Hz from a sensor placed on the belt buckle, the accuracy obtained in this model was 97.16%. In [10] off-the-shelf smartwatch was used in order to analyze the performance of ML algorithms. Using wrist-worn devices presents several challenges due to the positioning of the sensors, incorrectly placed sensors will produce more fluctuations in the data ads compared with sensors placed on the pelvis or the buckle. But by fusing multiple datasets (SmartWatch and SmartFall datasets), they were able to get an accuracy of almost 100%.

All the methodologies described in the background work use a single dataset or fused datasets collected from similar sensors. The split of the test and the training data is random and hence the same subject is used for training and testing the data. To our knowledge, no paper exists that uses separate datasets for testing and training.

3 Multiple Datasets

While working with privacy-sensitive data, especially healthcare-related, researchers face significant challenges in model development and debugging. Often the first step is to inspect individual examples in order to discover bugs, and outliers, generate hypotheses and improve labelling. In many cases, public datasets are used, direct inspection of data may be disallowed, and the data cannot be inspected. Public datasets become important if multiple datasets are used for training the model.

Several regulations, especially concerning healthcare data, allow the only collection of relevant data, and the data can be used only for the purpose it is collected for. The data collected cannot be used for future research. So, ideally, sensitive existing databases can only be used for different research directions without the possibility of privacy violations. Emerging from these problems and as a means to distribute the computational load of a training ML model, federated learning is proposed. The term FL was first used in the paper [11] and described a distributed and privacy-preserving way of training an ML model without others accessing private data. Even before federated learning can be applied, there is a need to study the effect of using multiple datasets on an ML model. The impact of training and testing with different datasets is still work to be done. To the best of our knowledge, literature is not available that studies the impact of multiple datasets in training and testing. The advantage of using multiple datasets would mean we can use data that is:

1. Massively distributed
2. Private
3. Unbalanced.

The datasets used in this paper, though, were collected using the same set of sensors but the method of collection and the demographics were hugely different. Some datasets have many data samples whereas some may have a limited number of samples available. Some of the challenges when using non-inspectable public datasets are:

1. Sanity checking
2. Model debugging
3. Data labelling
4. Detecting bias in training data.

3.1 Sanity Checking

Often the researchers will inspect some random examples and observe their properties before training a model. This is often done to identify the size, data type and range. This random check may also be used to identify outliers.

3.2 Model Debugging

When a model produces a result that is different from what is expected, it is natural to inspect a subset of input data, e.g. in a classification task, a modeller might inspect misclassified examples to look for issues in the features or the labels.

3.3 Data Labelling

For tasks where there is a large amount of data available, if low accuracy is observed on a specific slice of data, then it is possible especially in place of classification that the data is incorrectly labelled.

3.4 Detecting Bias in Trained Data

Each dataset might have a particular bias when multiple datasets are used. These biases must be checked for and corrected, especially when using a public dataset; it is impossible to know the data's process. The variation in the data collection process may add bias to the data. Hence, training on one dataset and testing on another will not produce the same accuracy as using the same dataset used for testing and training.

In this paper, we first use a single dataset with data collected from multiple individuals, data from some individuals are used for training and the rest are used for testing. We have used multiple datasets; while one dataset has been used for training, the others have been used for testing. By this methodology, we try to emulate the real-life scenario where data might be collected from multiple sources and used for testing. At the same time, the actual users of the model might be completely different individuals.

4 Data Collection Methodology and Public Datasets

4.1 Data Collection

We collected data using a TICWatch worn by ten volunteers; the TICWatch had multiple sensors such as a three-axes accelerometer, three-axes magnetometer, three-axes gyroscope, three-axes linear accelerometer and a five-axes rotation vector. Data was collected by asking the volunteers to perform 20 different ADL and fall activities. The ages of the volunteers were between 20 and 25 years, their height varied from 5.1 feet to 5.8 feet and their weight varied from 40 to 75 kg. Several of the volunteers had pre-existing health conditions such as malnutrition, claustrophobia, vertigo and diabetes. The presence of these medical conditions augurs well with naturally occurring health conditions that are a part of ageing. The following ADL activities were simulated by the volunteers: walking slowly and quickly, climbing up and down the stairs, jogging, transitioning from sitting to lying slowly and quickly, transitioning from a sideways position to lying back while still remaining in a lying position, standing up and sitting and getting up again, quickly sitting and standing up from a chair, stumbling, quick movements of the hands and jumping in place.

All directional falls were simulated (front, back, left and right) other than these grabbing while falling and spinning, while falling was also simulated. A programme was used to sense the data using the data and transmitting it via the user interface to a system where the data was stored. The data was moved by the system automatically into a csv file; buttons were provided on the user interface to perform every activity; as a result, labelled data was directly produced by the program.

For the safety of the user, these activities were simulated inside a well-padded anechoic chamber; the dataset generated had over a million data points of the ADL and the fall activities.

4.2 Public Datasets

Three public datasets were used, SmartWatch, Notch and SmartFall. The SmartWatch dataset [12] was collected from seven volunteers, each carrying an MS Band watch. The seven volunteers ranged from 21 to 55; height ranged from 5 to 6.5 ft and weight from 45 to 104.5 kg. Each volunteer performed a pre-determined set of ADLs: jogging, sitting down, throwing an object and waving their hands. And then, the volunteers were asked to fall on a 12-inch mattress on the floor, namely front, back, left, and right falls.

Using a wrist-worn Notch sensor [13], the Notch dataset was collected from simulated fall data and ADLs. The Notch system consists of multiple individual sensors placed on different parts of the human body to collect motion data. Seven volunteers with ages ranging from 20 to 35 and heights from 5 to 6 ft, and the weight varied from 45 to 90 kg, were used. The Notch sensor was paired to an android device (tablet) via Bluetooth through a custom-built data collection app. A list of seven ADLs, sitting up, getting up, jogging, throwing an object, waving, taking a drink and going up and down the stairs and four types of falls, front, back, left, and right, were performed by the volunteers.

The SmartFall datasets [14] used 14 subjects covering a wide age range of 21 to 60 and 1027 activities and 92,780 data points.

4.3 Feature Extraction

Statistical data were collected from all four data sets. The statistical parameters collected were mean, standard deviation, variance, minimum, maximum, skew and kurtosis. The ML algorithms were run using the statistical data as input.

We can observe from the four data sets that the volunteers used the actions performed and the method of collecting data are entirely different. Hence, we can better analyze the ML algorithms' accuracy by using multiple datasets and their statistical parameters.

5 Results and Discussions

5.1 BITS Dataset

To understand the impact of training on a different dataset and testing on a different dataset, we initially used only the BITS dataset and split the volunteers into two sets. A total of 70% of them were used for generating the training data while 30% of them were used for testing. We also ran the ML algorithms using the traditional method of randomly splitting the data. The results of this are shown in Fig. 1.

From Fig. 1, it can be seen that when we use a separate dataset for testing and training, there is a drop in accuracy in the case of all the four ML algorithms that is (a) KNN [15] (b) logistic regression [16] (c) Naive Bayes [17] (d) random forest [18].

The impact of accuracies in random forest is minimal as it employs multiple decision trees that progressively learn from each other. In the case of KNN, there is a drop in accuracy of about 15%; in the case of logistic regression is 18% and in the case of Naïve Bayes 26% while in the case of random forest the drop is only 5%. This variation, as can be observed from Fig. 1 indicates that when machine learning algorithms use separate training and testing datasets even while the demographics, the sensors and the data collection methodology are similar, there is a drop in accuracy. This indicates that with federated learning the performance of the machine learning algorithms will improve.

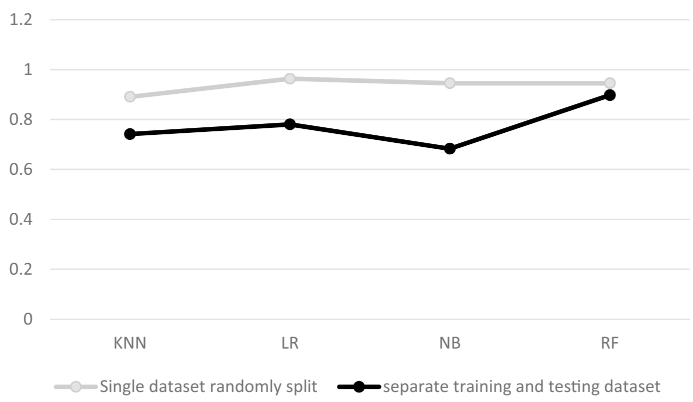


Fig. 1 Randomly split dataset versus separate dataset for training and testing

5.2 Public Datasets Combinations

To further explore this, we used three public datasets; Notch, SmartFall and SmartWatch datasets. We tried various combinations of the three datasets for testing and training.

Notch Dataset

We tried the following testing and training combinations assuming that the Notch dataset will be the final user.

1. Training and testing with Notch
2. Training with SmartFall and testing with Notch
3. Training with SmartWatch and testing with Notch.

The results are depicted in Figs. 2 and 3.

It can be seen very clearly from Figs. 2 and 3, when using different datasets for training and testing there is a big drop in accuracy. From Sect. 3, where the data

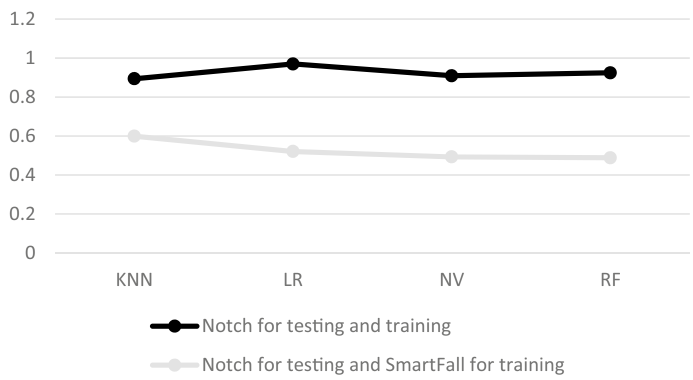


Fig. 2 Notch for testing and training versus SmartFall for training

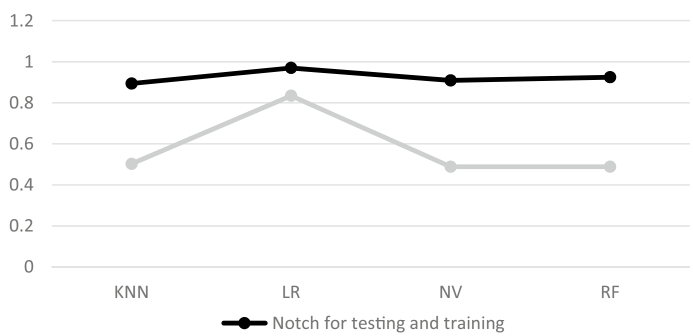


Fig. 3 Notch for testing and training versus SmartWatch for training

collection methodology is described, it can be observed that even though similar sensors were used in the demographics of the volunteers, the ADL activities and the fall activities were different, which is close to what we can get in real-life scenarios as neither the ADL nor the Fall will be planned activities for a user. Hence, the prediction may be completely incorrect. In the case of testing and training with Notch, the accuracy varies from 90 to 96%, respectively, but when SmartFall was used for training, there is a drop in accuracy that varied between 30 and 48%. The maximum drop in accuracy was observed in random forest even though progressive learning was used. KNN in this case gave the best results, the drop in accuracy being only 30% which is still a large drop. As the datasets do not describe clearly the methodology used for collecting the data, neither do they give the details of which data is matched with which volunteer due to data privacy issues, it is difficult to analyze why KNN seems to have a lesser drop when compared with random forest.

SmartWatch Datasets

We tried the following testing and training combinations assuming that the Notch dataset will be the final user.

1. Training and testing with SmartWatch
2. Training with Notch and testing with SmartWatch
3. Training with SmartFall and testing with SmartWatch.

The results are depicted in Figs. 4 and 5.

It can be observed from Fig. 4, when the Notch is used for training with SmartWatch for testing, there is a huge drop in accuracy which varies from 28% in logistical regression to 67% in the case of random forest. This huge drop in random forest can be explained by the fact that the number of data points in the case of Notch (10,645) is much lesser than the number of data points that could be extracted from SmartWatch (34,019).

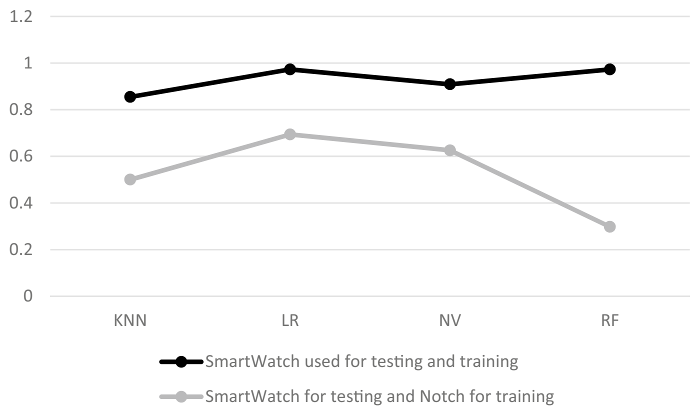


Fig. 4 SmartWatch for testing and training versus Notch for training

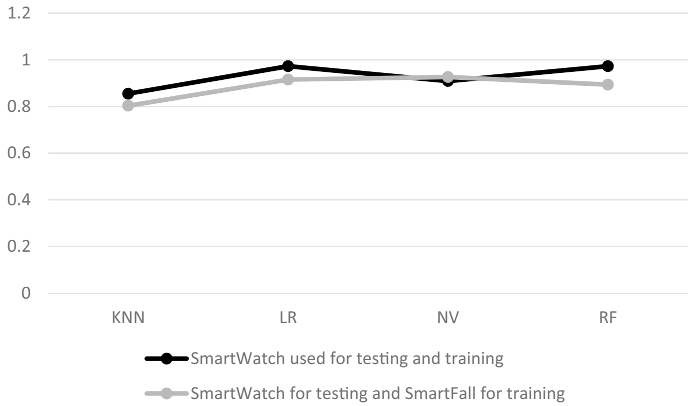


Fig. 5 SmartWatch for testing and training versus SmartFall for training

In the case of Fig. 5, it can be observed that the drop in accuracy is almost negligible; in fact, in the case of Naïve Bayes the drop in accuracy is almost 0%; this is because SmartFall and SmartWatch, other than using the same sensors used volunteers with the same demographics. The data collection methodology cannot be commented upon as the details were not provided in the public dataset. In reality, it is difficult to train the system using volunteers from the geriatric age range while people above 50 or 60 can be asked to simulate daily activities; it is risky to subject them to simulated falls. SmartFall and SmartWatch used users in the geriatric age range to simulate daily activities, hence, the ML algorithms yield a similar accuracy.

SmartFall Datasets

We tried the following testing and training combinations assuming that the Notch dataset will be the final user.

1. Training and testing with SmartFall
2. Training with Notch and testing with SmartFall
3. Training with SmartWatch and testing with SmartFall.

The results are depicted in Figs. 6 and 7.

It can be observed from Fig. 6 that the pattern of drop in accuracy between 28% in the case of logistic regression to 70% in the case of random forest follows a similar pattern to Fig. 4. The reason again is that the number of data points in the case of Notch (10,645) is much lesser than that of SmartFall (92,780) and the number of volunteers was twice that of Notch. Also, the SmartFall demographics were completely different from that of Notch. Hence, the variation in accuracy and especially the drop in the case of random forest is expected.

Again, the similarities in accuracies from using SmartWatch for training and SmartFall for testing are expected as can be observed from Fig. 7. The accuracies are similar since the demographics, sensors and data collection methodologies remain the same for SmartWatch and SmartFall.

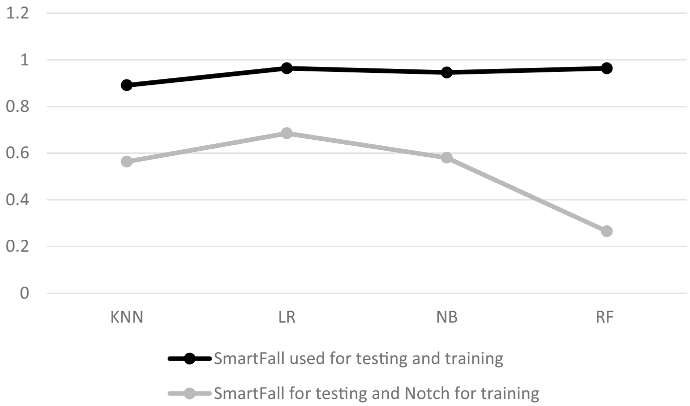


Fig. 6 SmartFall for testing and training versus Notch for training

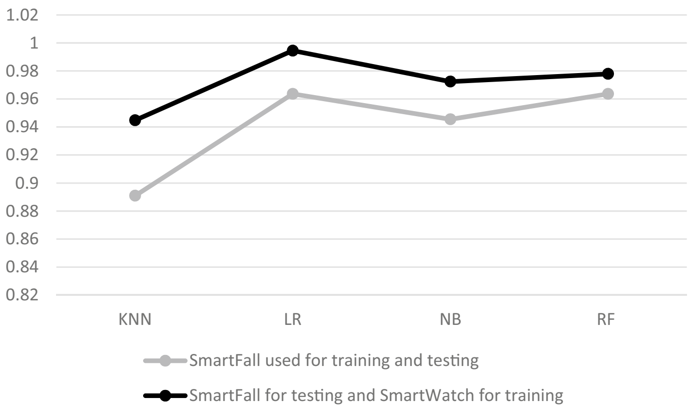


Fig. 7 SmartFall for testing and training versus SmartWatch for training

6 Conclusion

In a data-centred world where people are expected to share their data, it is very important to preserve data privacy especially when it is regarding health-sensitive information. It is also necessary that a large amount of data is required for training ML and DL algorithms. In the case of an application such as fall detection, not many public datasets with multiple volunteers are available. To the best of our knowledge, the number of volunteers has not exceeded 14–15 also if we are talking about fall detection in geriatrics, the demography of the training volunteers is expected to be completely different from the demographics of the end users. Though multiple smart devices such as smartphones and smartwatches have been used to generate public datasets, they have certain shortcomings in terms of the kind of sensors used;

these sensors essentially are not medical-grade sensors but rather are used to track certain fitness activities. At the same time, the volunteers used for collecting the data are usually in the age range of 20–40. As it is a huge health risk to ask anyone above 40 years to fall. Also, several of the public datasets available simulate very few ADL activities and the ADL activities simulated have very less resemblance to the fall activities. Hence, the extracted data gives high accuracy. In real-world scenarios, many daily activities may resemble fall activities, especially kneeling on the ground, lying on the ground, getting up/sitting up abruptly, bending, etc. Hence, it is possible that some of the daily activities may be misinterpreted as fall activities. The solution to this problem is to train the device using multiple datasets collected from wide demographics in terms of nationality, age, gender and pre-existing health conditions. A single dataset may not suffice to train the system, although attempts have been made to fuse the datasets []; the fused datasets are from public datasets that use similar sensors, similar demographics, similar activities, and similar data collection methodologies. The solution to this is to use a widely distributed dataset for incremental training of the various ML and DL algorithms. This leads us to the necessity of using federated learning. Federated learning has only gained prominence in the last couple of years; while there are studies of federated learning techniques, the actual implementation of fall detection is yet to be carried out. In this paper, we have experimented with our dataset (BITS dataset) as well as public datasets to understand the impact of using multiple datasets for testing and training. From our results, it is obvious when the test and the train demographics are completely different, and an insufficient number of activities are simulated there is a huge drop in accuracy. The only way forward especially in the area of geriatric fall detection is to use federated learning across a distributed dataset sourced from multiple public datasets and clinical data. Our future work will concentrate on implementing federated learning with the use of incremental learning in decision trees to develop a proper ML/DL model that can be used for medical grade fall detection applications.

References

1. World Health Organization (WHO) Ageing and Health. Fact Sheet. Available online: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health
2. Casilari E, Lora-Rivera R, García-Lagos F (2020) A study on the application of convolutional neural networks to fall detection evaluated with multiple public datasets. Sensors 20:1466. https://doi.org/10.3390/s20051466
3. Ramachandran A, Karuppiah A (2020) A survey on recent advances in wearable fall detection systems. BioMed Res Int 2020:17. Article ID 2167160. https://doi.org/10.1155/2020/2167160
4. Habib MA, Mohktar MS, Kamaruzzaman SB, Lim KS, Pin TM, Ibrahim F (2014) Smartphone-based solutions for fall detection and prevention: challenges and open issues. Sensors 14:7181–7208
5. Bourke AK, Lyons GM (2008) A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. Med Eng Phys 30:84–90

6. Theodoridis T, Solachidis V, Vretos N, Daras P (2018) Human fall detection from acceleration measurements using a recurrent neural network. In: Precision medicine powered by pHealth and connected health; Springer, Singapore, pp 145–149
7. Kwolek B, Kepski M (2014) Human fall detection on embedded platform using depth maps and wireless accelerometer. Comput. Meth Prog Biomed 117:489–501
8. Musci M, De Martini D, Blago N, Facchinetti T, Piastra M (2018) Online fall detection using recurrent neural networks. ArXiv 2018, arXiv:1804.04976
9. Sucerquia A, López JD, Vargas-Bonilla JF (2017) SisFall: a fall and movement dataset. Sensors 17:198
10. Mauldin TR, Canby ME, Metsis V, Ngu AHH, Rivera CC (2018) SmartFall: a smartwatch-based fall detection system using deep learning. Sensors (Basel). 18(10):3363. doi: https://doi.org/10.3390/s18103363.PMID:30304768;PMCID:PMC6210545
11. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA (2017) Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th international conference on artificial intelligence and statistics, in proceedings of machine learning research vol 54, pp 1273–1282. Available from https://proceedings.mlr.press/v54/mcmahan17a.html
12. SmartWatch dataset available online: https://userweb.cs.txstate.edu/~hn12/data/SmartFallDataSet/SmartWatch/
13. Notch dataset available online: https://userweb.cs.txstate.edu/~hn12/data/SmartFallDataSet/notch/Notch_Dataset_Wrist/
14. SmartFall dataset available online: https://userweb.cs.txstate.edu/~hn12/data/SmartFallDataSet/SmartFall/
15. Alpaydin E (1997) Voting over multiple condensed nearest neighbors. Artif Intell Rev 115–132
16. Hosmer DW, Lemeshow SL (2000) Applied logistic regression, 2nd edn. Wiley-Interscience, Hoboken, NJ
17. Vembandasamy K, Sasipriya R, Deepa E (2015) Heart diseases detection using naive bayes algorithm. IJISET—Int J Innovat Sci Eng Technol 2(9). ISSN 2348-7968
18. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees (1st ed.). Routledge. https://doi.org/10.1201/9781315139470

Is Internet Language a Destroyer to Communication?

Chan Eang Teng and Tang Mui Joo

Abstract Internet language is known as the new form of language that has been used on social media by the Internet users. Since the Internet language has been widely used through social media, there is some influence on the users' behaviour. As a result, the use of the Internet language is feared to undermine the authenticity of the original language. Language is a significantly important communication tool for everyone, and hence, there are many studies on human communication. As the emergence of the Internet is growing fast, language has been affected and caused by the inventions of Internet language which is also known as Internet slang. It is also now widely used by people in their daily communication. However, there are several problems caused by Internet language. For example, people who seldom use the Internet could not understand the Internet language which might cause communication problems, the loss of language authenticity, and generation gap. As the study of Internet language is not broad, a few problems still remain unknown such as the communication habits of Internet language users, the level of understanding of the original language, and how elders are out of touch with contemporary society due to the Internet language. Quantitative research method which is an online survey is used to study the generation Z who were born from 1997 to 2012 and baby boomers who were born from 1955 to 1964. The reason why this research targets these selected samples is to investigate the generation gap between gen Z and baby boomers that the Internet language brings to them. The research is intended to investigate how the Internet language affects human communication habits. This research has found out that the Internet language has actually affected human communication habits, because the Internet language has become a part and parcel of their communication style.

Keywords Internet communication · Internet language · Internet slang · Communication · Language

C. E. Teng (✉) · T. M. Joo
Tunku Abdul Rahman University of Management and Technology, 53300 Kuala Lumpur, Malaysia
e-mail: eangteng@hotmail.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_42

1 Introduction

Internet language is widely used when communicating with each other. Nevertheless, it has drastically changed the ability of understanding certain terms or words that are used within the networking or in daily conversation among the users. For example, one of the earlier Internet languages used on social media is "LOL", laugh out loud to indicate a funny thing by Wayne Pearson back in the 1980s when he was chatting with his friend [1]. LOL now has been inducted into the Oxford English Dictionary [2]. As of now, many more Internet languages are flooding the social network. The Internet language is able to express the message in a more interesting way and could have saved the message typing time by using an abbreviation or short form. No doubt, every coin has two sides, and effects might occur as it shows. The "new vocabulary" might confuse other people who might not be familiar with the meaning and cause misunderstandings. Hence, this study is to identify the impact of Internet language on social media on communication behaviour among people. The area of the research included the communication style, the impacts on language itself, and impacts on the older generation.

2 Literature Review

2.1 Internet Slang and Communication Style

Communication style is basically the method of an individual interacting and exchanging information with one another [3]. Due to the rapid growth of the Internet and social media, it causes the development of a new language called Internet language, Internet slang, and texting slang. A common question raised is, why do people in contemporary society like to use Internet language in their daily communication? As Azida Sabri [4] found out, the use of Internet language could be associated with three factors that are secrecy, time, and trend. Sometimes people use Internet slang all the way in their communication and sometimes only a little [5]. By using Internet language, communication between one another turns to be more secretive [5]. This will develop a mysterious communication style where the words, phrases, and even the whole conversation might not be completely understood by someone who does not join the conversation.

2.2 Internet Language Destroys the Authenticity of Language

Texting and communication online have emerged as the mainstream of young people that predominantly use non-standard language or can be understood as Internet language [6]. Majority of the young people who use social media like to apply homophones and abbreviations [7]. As a result, the Internet language has affected the vocabulary of the language which will cause the youth that have not shaped their self-consciousness to follow and gradually lead to lost authenticity of language. Apart from that, not only youth but also the elderly is influenced by the Internet from time to time. People are unaware that they have been communicating in a manner that is different from their typical daily conversation. The depth and quantity of information available on the Internet gave way to a new language known as Internet language [8]. Besides, people who are non-Internet users need to take a longer time to understand it. When it comes to Internet language, it has been a significant component of the Internet itself since it first appeared, and it also becomes more difficult to understand that more often used as time goes on [8]. Here, the question raised, whether the Internet language will deteriorate and challenge the authenticity of the original language.

2.3 The Impact of Internet Language on the Older Generation

The use of Internet language and new words is becoming more and more contagious among teenagers due to the phenomenon of language modernization among generations Teenagers today can easily comprehend and use the Internet slang for their community [9]. Everyone shall witness the emergence of a new culture, which is inextricably linked to the emergence of a new language [10]. The older generation has been misunderstood in social network conversations because of the non-standard language used in online communication. Older generations are slower to grasp the digital world, more cautious, and less receptive to the changes that are occurring because they did not grow up with this modern phenomenon [11]. According to Eliza and Marigrace [12], language changes across space and social groups. Most of the words that the older generation do not understand come from online social platforms when the young generation gathers. Generation Z individuals need to formalize their writing and refrain from using superfluous spelling innovations, capitalization, abbreviations, emotions, and punctuation that may confuse baby boomers in order to minimize discrepancies and close the communication gap [11]. It is also possible to direct the development of the communication system in favour of the ideas of a just society [13]. The older generation should also be knowledgeable in online literacy in order to interact with the younger generations [11].

3 Theoretical Framework

Informalization is a theory that states that informal language is now being formalized and used in society instead of within close relationships. Internet language is a very informal language due to its unusualness like abbreviation and short form. Informal style of language has brought some influence to individuals and society, especially the communication habits and style. This is because informal language like the Internet language leads a conversation towards a more relaxed manner which will help in smoother communication among people. However, using it in formal situations like business meetings will show unprofessionalism. Random Fluctuation theory explains that language is unstable and changes may often occur unpredictable. Language is not fixed but is constantly changing based on the culture, situations, and behaviour of an individual or society. When those multilingual people use the languages they are experts in, they might bring a new vocabulary or use it on social media platforms. At the same time, those who do not understand the meaning behind might interpret the languages via their own understanding and more homophonic words are created.

4 Methodology

4.1 Subjects/Participants

Quantitative method is employed to carry out the research where online survey is chosen as it is the most common method that is widely used to collect data from a group of people by asking them questions when it needs large sample size to generalize the data. A total of 107 samples of Generation Z (1997–2012) and baby boomers (1955–1964) are included in this online survey. The target population was chosen because generation Z is more accustomed to the internet and new media than baby boomers, who have less exposure to and knowledge of this Internet and also new media. Therefore, it is able to let the research continue about the communication behaviour between Internet language users (Gen Z) and non-users (baby boomers). Besides that, from the target population, the research can also get opinions about the topic of “how far the Internet language has destroyed the authenticity of language”. Furthermore, this research intends to show how seriously the Internet language has impacted on the older generation which is the baby boomers by targeting generation Z and baby boomers to fill up the online google survey form.

4.2 Research Design

The research method that is used in this research is the quantitative research method, which is an online survey questionnaire. The survey questionnaire includes three main sections which are demographic, social media behaviour, and the impact of Internet language on communication habits. The online survey is conducted from 5 August 2022 to 25 August 2022 using convenience sampling.

5 Result

5.1 Demographic Profile

As shown in Table 1, the majority of the respondents are occupied by the age group of 18–24 which have 59.8%, 64 respondents out of 107 respondents. Followed by respondents aged 50 and above which is 13.1%. Meanwhile, the lowest percentage which is 2.8% are under 18 respondents. The respondents are highly participated by females compared to male respondents which are 69.2% and 30.8% accordingly. As for the occupation status of this survey participated mostly by students which are 57.9%, while the lowest percentage goes to the self-employed and housewife or househusband, 9.3% as both of them have the same percentage.

Table 1 Demographic profile

Age	Frequency	Valid percentage (%)
Under 18	3	2.8
18–24	64	59.8
25–34	13	12.1
35–49	13	12.1
50 and above	14	13.1
Gender	Frequency	Valid percentage (%)
Male	33	30.8
Female	74	69.2
Occupation	Frequency	Valid percentage (%)
Student	62	57.9
Employee	25	23.4
Self-employed	10	9.3
Housewife/ Househusband	10	9.3

Source Online survey conducted from 5 to 25 August 2022

Table 2 Encounterment of Internet language

How do you encounter Internet language?	Frequency	Valid percentage (%)
Through conversation with others	71	66.4
Internet and social media	97	90.65
Magazine/newspaper	10	9.3
TV series/drama	31	29

Source Online survey conducted from 5 to 25 August 2022

Table 3 Level of understanding of Internet language

What is your level of understanding of Internet language?	Frequency	Valid percentage (%)
0	5	4.7
1	7	6.5
2	19	17.8
3	28	26.2
4	39	36.4
5	9	8.4

Source Online survey conducted from 5 to 25 August 2022

5.2 Understanding of Internet Language

Table 2 shows the majority of the respondents encountered Internet language via Internet and social media where Internet and social media have been the main source of Internet language. Table 3 shows the majority of the respondents possess above average level of understanding of Internet language. This is in line with the finding of Fish [9] that teenagers today can easily comprehend and use the Internet language since they are active on social media. Table 4 indicates most of the respondents use Internet language in their daily communication be it online or offline. This is due to the features provided by Internet language such as convenient, entertaining, expressing emotions, and following the trend as shown in Table 5. Based on Rezeki and Sagala [14], it explains that Internet slang could make conversations more relaxed and comfortable. Table 4 also shows the usage of Internet language has become very common in all types of communication but not limited to only online daily communication. However, the result shows the usage of Internet language in academics is somehow limited.

5.3 Impact of Internet Language on Communication Habits

Tables 6 and 7 show the impact of Internet language towards the respondents. Majority of them acknowledge that the heavy use of Internet language might ruin

Table 4 The usage of Internet language

When do you use Internet language?	Frequency	Valid percentage (%)
Daily communication with others	53	49.5
Online communication with others	77	72
Phone texting	69	64.5
In academic	8	7.5
I don't use Internet language	19	17.8

Source Online survey conducted from 5 to 25 August 2022

Table 5 Reason of using Internet language

Why do you think people use Internet language?	Frequency	Valid percentage (%)
Convenient	70	65.4
Entertainment	68	63.6
Emotions express	68	63.6
Follow the trends	56	52.3

Source Online survey conducted from 5 to 25 August 2022

or destroy the grammar and authenticity of the language. At times it also creates misunderstanding as there is no standardization of the interpretation of internet language and inability to type or interpret long original messages. However, the usage of Internet language is less likely to make people less socialized. Table 7 also shows that Internet language makes the communication funny and humorous. It helps to improve the communication speed, gives space to cultivate creativity, eases the process of learning language, and builds some bonding in the community.

From the correlation coefficient test, it shows that the correlation between the *Tendency of Using Internet Language in Daily Communication* and *Entertainment* is highly significant. The higher the tendency of people to use Internet language in their daily communication, the more entertaining it is. This is in line with the review about how the Internet language will have an impact on the communication style

Table 6 Negative impact of Internet language

Negative impact of Internet language	Frequency	Valid percentage
Ruin or destroy the grammar and authenticity of language	76	71
Misunderstanding	71	66.4
Make people become lazy in typing long messages	57	53.3
Detrimental to language development	37	34.6
Decrease the ability to perceive and use original language	47	43.9
Make people less socialize	8	7.5

Source Online survey conducted from 5 to 25 August 2022

Table 7 Positive impacts of Internet language

Positive impact of internet language	Frequency	Valid percentage
Improve the communication speed	72	67.3
Able to cultivate people's language creativity	56	52.3
Communication among people becomes funny and humorous	82	76.6
Make it easier for people to understand and learn language	37	34.6
Symbolizes community identity	33	30.8
Building stronger bonding in community	39	27.1

Source Online survey conducted from 5 to 25 August 2022

of an individual as it makes the conversation more relaxed and comfortable [14]. This also implies that the Internet language in a conversation is more entertaining compared to the original language. The correlation coefficient test also indicates the correlation between *Age* and *The Level of Understanding of Internet Language* is highly significant. The older the person is, the lower their understanding of the Internet language. This is also in line with the review that the older generation could not understand and sometimes misunderstand the Internet language that frequently appears in social networks compared to teenagers [11]. Another correlation between *Daily Communication with Internet Language* and *The Ability to Perceive Original Language* is highly insignificant. The more people use Internet language in their daily communication, the more their ability to perceive and use original language will decrease. This is in line with the review that the Internet language has affected the vocabulary of the original language [9]. The test also shows the correlation between *Daily communication* and *The Cultivation of Language Creativity* is significant. When more people use the Internet language in their daily communication, the more it will cultivate their language creativity. The use of Internet language can cultivate people's language creativity by creating new vocabularies which have transformed into a new form of language. This finding has challenged the previous finding which indicates that the Internet language will destroy the authenticity of language.

6 Conclusion and Discussion

As social media is now flooded with different kinds of Internet language, people are forced to understand the meaning of Internet slang in social networks not only in daily communication but also in the media content. Informalization theory stresses that informal language is now being formalized and used in society instead of within the close relationship. The result also implies the instability of one language as it might change at times as explained in random fluctuation theory. This is shown when some of the Internet languages which used to be informal are now included in the Oxford Dictionary as the formal language. And at the same time, the Internet language also changes from time to time depending on the usage and popularity of it. This is in line with the highlight of random fluctuation theory that languages might be insatiable and fluctuated. The result also implies that Internet language makes the communication more relaxing and entertaining which has added some elements in communication. With the speed that saves more time in communication, Internet language is a bonus in human communication. The Internet language has obviously affected the communication habits of Internet users as they will use it in online communication and daily communication. It is also worth mentioning that the unique cultural phenomenon of the Internet language not only provides people with a language form for communication, but also opens up new research space for other areas of communication such as organizational communication, corporate communication, or communication in academics. However, some still hold a comparatively negative attitude towards the Internet language as it is not standardized, and this might create misunderstandings in communication. Many opined that Internet language is convenient and entertaining as it provides a sense of humour and makes people feel more relaxed in communication. On the other hand, people also have the fear that Internet language will ruin the authenticity of language and cause misunderstanding in daily communication. Due to the time complexity, this research only reports the generalized data collected from the online survey to gauge the idea of how the Internet language has impacted both the baby boomers and Generation Z which focuses on their daily communication. The result of this research shall be further discussed by investigating the impact of Internet language in other forms of communication.

Acknowledgements The authors acknowledged the raw materials provided by Alice Tan, Hor Yan, Wern Jing, and Sherwyn Yap.

References

1. Unbabael (2019) Do you speak internet? How internet slang is changing language. https://resources.unbabel.com/blog/speak-internet-slang. Last accessed 5 Aug 2022
2. Kern R (2015) Language, literacy and technology. Cambridge University Press, Cambridge
3. Alvernia University (2000) Applying business models to higher education. https://academicjournals.org/journal/IJEAPS/article-full-text-pdf/1380FAC58982. Last accessed 22 Sep 2022
4. Sabri NAB, Hamdan SB, Nadarajan N-T M, Sing SR (2020) The usage of English internet slang among malaysians in social media. Selangor Humaniora Review
5. Coleman J (2022) The life of slang: the history of slang. https://ebookcentral-proquest-com.tarcez.tarc.edu.my/lib/tarc-ebooks/reader.action?docID=943382. Last accessed 23 Aug 2022
6. Farina F, Lyddy F (2011) The language of text messaging. “Lingustic ruin or resource?” The Irish Psychologist 37(6):145–149
7. Kadir ZA, Idris H, Husain SSS (2012) Playfulness and creativity: a look at language use online in Malaysia. Proced-Soc Behav Sci 65:404–409
8. Indera WAIWA, Ali AAER (2021) The relationship between internet slang and English language learning 4(2):1–6
9. Fish TW (2015) Internet slang and high school students: a teacher’s perspective. A thesis submitted for Master of Arts in Communication and Leadership Studies, Gonzaga University
10. Petrova YA, Vasichkina ON (2021) The impact of the development of information technology tools of communication on digital culture and Internet. https://doi.org/10.1051/shsconf/20110101002. Last accessed 23 Aug 2021

11. Subramaniam V, Razak NA (2014) Examining language usage and patterns in online conversation: communication gap among generation y and baby boomers 118:468–474
12. Eliza MJ, Marigrace DC (2022) Digital culture and social media slang of gen Z, https://uijrt.com/articles/v3/i4/UIJRTV3I40002.pdf. Last accessed 5 Aug 2022
13. Mansell R (2021) Imagining the internet. https://ebookcentral-proquest-com.tarcez.tarc.edu.my/lib/tarc-ebooks/detail.action?docID=998950&pq-origsite=summon. Last accessed 2021/08/20. Last accessed 20 Aug 2021
14. Rezeki TI, Wahyudin R (2019) Language Acquisition Pada Anak Periode Lingustik. Serunal Jurnal Ilmiah Ilmu Pendidikan 5(1):84–89

Variational Autoencoders Versus Recurrent Neural Network for Detection of Anomalous Trajectories

Muhammad Ehsan Siddique, Yousra Chabchoub, Michele-Luca Puzzo, and Ammar Kheirbek

Abstract Anomalous trajectory detection has a prominent place in many real-world applications, as for example taxi fraud detection. Several approaches have been proposed over time, but in this paper the goal is to analyze and compare two novel models (GM-VSAE and ATD-RNN) which have addressed the problem differently. However, both of them have already overcome traditional anomalous trajectory detection methods. For this purpose, we have worked on a real-world taxi trajectory dataset, in which we introduced some anomalies. First, we conducted an explanatory analysis of this dataset. Then, we explained the principles of the two models, highlighting their differences and finally, we evaluated their performances on the considered dataset. Results show that GM-VSAE is more efficient even if both models have shown their relevance in detecting anomalous trajectories.

Keywords GM-VSAE · ATD-RNN · Anomalies detection · Trajectories

M. E. Siddique (✉) · Y. Chabchoub · A. Kheirbek
ISEP—Institut Supérieur d'Électronique de Paris, 10 rue de Vanves, Issy les Moulineaux, 92130 Paris, France
e-mail: ehsan.siddique@cri-paris.org

Y. Chabchoub
e-mail: yousra.chabchoub@isep.fr

A. Kheirbek
e-mail: ammar.kheirbek@isep.fr

M. E. Siddique
Université Paris Cité, 8 bis Rue Charles V, 75004 Paris, France

M.-L. Puzzo
University of Rome (Sapienza), Piazzale Aldo Moro, 5, 00185 Roma, RM, Italie
e-mail: puzzo.1783133@studenti.uniroma1.it

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_43

1 Introduction

The advancement in technology has increased spatio-temporal data. Different techniques can be used for data-mining trajectory data. These techniques are used in different stages like pre-processing, data management and other variety of tasks (such as trajectory pattern mining, outlier detection and trajectory classification). In the pre-processing stage, we have to deal with several issues, such as noise filtering, segmentation, map matching, trajectory compression and stay points. Data management stage deals with the trajectory segmentation and divides the trajectory based on the time interval, the spatial shape and the semantic meaning before both classification and clustering.

We can also transform the trajectories into different forms like graphs, matrices or tensors. After transformation, we can perform another series of steps like collaborative filtering, matrix factorization and tensor decomposition [1]. Trajectories can be divided into four categories: mobility of people, mobility of transportation vehicles, mobility of animals, mobility of natural phenomena. Trajectory outlier detection has become an important task in the trajectory analysis. Most of the traditional trajectory detection algorithms are based on classification, clustering and distance or density-based statistics.

These algorithms use different distance metrics for the detection of outliers, such as Euclidean, Hausdorff, Longest Common Subsequence (LCSS) [2]. These traditional methods cannot handle the variety and the complexity of trajectory data and do not provide efficient online trajectory outlier detection [3]. Deep learning methods can solve this problem while being more efficient than the traditional methods. Due to the sequential nature of the trajectory data, we have compared in this paper two sequential methods proposed by Liu et al. [3], Song et al. [4]. We have also tested these methods on real taxi trajectories dataset: Taxi Service Trajectory. Finally, we have evaluated the performance of the models based on the architecture and evaluation metrics, i.e., precision, recall and F1-score.

2 Related Work

Data mining has become very important in the analysis of trajectories, because of the complexity and variety of trajectory data. The main tasks of trajectory mining can be categorized into four categories: trajectory classification, trajectory prediction, trajectory pattern mining, outlier detection. We also need to deal with several issues, such as noise filtering, segmentation and map matching during the pre-processing stage. Noise filtering is done generally to improve data quality, due to the poor signal of the GPS device. Different filters can also be used to remove noise in spatial data. These filters include mean or median filter, Kalman and particle filter and heuristic-based outlier detection.

Table 1 Summary of distance metrics

Metric	Literature year	Time complexity	Anti-noise property
Euclidean	Clarke [7]	$O(n)$	Weakest
Hausdroff	Huttenlocher et al. [8]	$O(m * n)$	Weak
Longest common sub-sequence (LCSS)	Robinson [9]	$O(m * n)$	Strong
Dynamic time wrapping (DTW)	Sankoff and Kruskal [10]	$O(m * n)$	Weak

Stay point is the point where the moving object has stopped for a certain time [1]. Sometimes spatial data is not of the same importance because of the stay points. To solve this problem, Li et al. [5] proposed a stay point detection algorithm. It checks if the distance between the anchor and the last successor point is larger than the given threshold. It then measures the time between the anchor and the successor, so the measured time is larger than the threshold and the point is considered to be a stay point.

Trajectories segmentation is important, because most of the trajectories are complex, and it is difficult to extract the pattern for the classification or clustering. Segmentation of trajectory can be done using the time interval, and the shape of the trajectory, and also by the semantic meaning [1]. We can use the Douglas-Peucker algorithm to find out the key points of the trajectory for the simplification of the trajectory [6].

The detection of anomaly in trajectory is a very important task too in the analysis of the trajectory data. Different methods are used for the detection of anomalies, such as classification-based methods, clustering-based methods, distance-based methods, density-based methods and statistic-based methods [2]. Clustering methods can be effective, as most of the trajectory data is not labeled.

Distance-based methods find out the outlier by calculating the distance between two objects. One of the best techniques to find out the distance is to use the k-nearest neighbors. The worst time complexity of this algorithm is $O(n^2)$, where n is the number of trajectories. Different distance metrics can be used for the distance-based methods. Density-based methods depend on distance-based methods because the density is usually defined by the Euclidean distance. The worst-case time complexity is the same in the distance-based methods. The distance-based and density-based methods are not effective approaches if the data is large. The summary of the distance-based metrics is given in Table 1.

Statistical-based methods incorporate probabilistic models and perform well for large amount of data, but they are not so ideal for the high-dimensional data. Hidden Markov Model (HMM) is used for the high-dimensional data. Suzuki et al. [11] used HMM to model the spatial and temporal features of human trajectories in video data. They used eigenvector decomposition to project the data from high-dimensional space to the low-dimensional space. The outlier can be detected using the likelihood score of the HMM.

Lee et al. [12] introduced a partition-and-detect framework and presented an Abnormal Trajectories Outlier Detection (TRAOD) algorithm to detect the anomalies, using the shape of dissimilarities and the local motion of the sub-trajectory. The distance metric used in this method is hausdroff distance which does not involve the common deviation between the sub-trajectories. This distance-based methods require suitable parameters. To overcome this problem, Liu et al. [13] introduced the Density-Based Trajectory Outlier Detection (DBTOD) algorithm, which can detect both anomalous sub-trajectories and anomalous local trajectories. Both of these methods used partition and detection framework.

3 Dataset Description

We considered the public dataset *Taxi Service Trajectory*[1] which describes one-year trajectories performed by 442 taxis running in the city of Porto, in Portugal, from 01/07/2013 to 01/07/2014. It contains 1,710,670 data samples and each one corresponds to a completed trip. Each trip is characterized by 9 features. For each trip the GPS coordinates of the taxi are given every 15 seconds. We focused on the following important 5 features:

- **TRIP_ID**: a unique identifier for each trip.
- **TAXI_ID**: a unique identifier for the taxi driver who performed each trip.
- **TIMESTAMP**: the start time of the trip.
- **MISSING_DATA**: it is *False* when the GPS data stream is complete and *True* whenever, at least, one location is missing.
- **POLYLINE**: trajectory of the trip in the form of a list of GPS coordinates (WGS84 format), in which each pair [longitude, latitude] is taken each 15 seconds of the trip.

3.1 Explanatory Data Analysis

First we have preprocessed our dataset removing trips in which the GPS data stream is not complete (*MISSING_DATA* = *True*), and ones in which the trajectory is constituted by just zero or one pair of coordinates.

We plotted in Fig. 1 the number of trips or trajectories performed per month. In average, 140,000 trajectories are performed each month. The month of May 2013 corresponds to the highest number of trajectories.

[1] https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i/data.

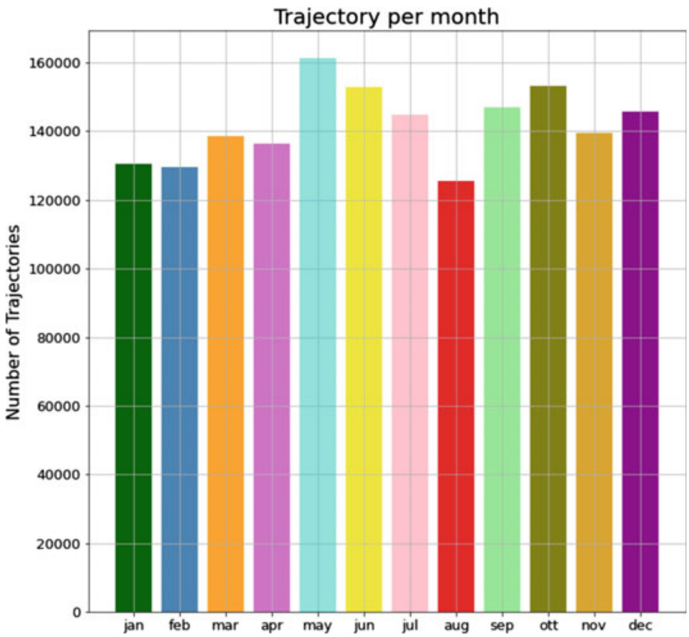


Fig. 1 Number of trajectories per month

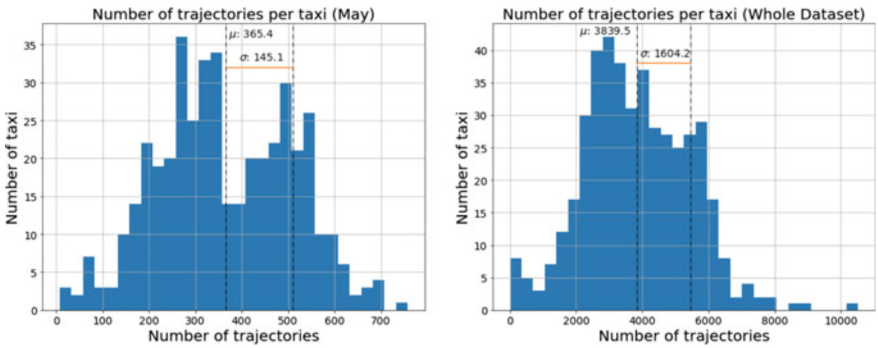


Fig. 2 Distribution of number of trajectories per taxi

Figure 2 represents the distribution of the number of trajectories per taxi, respectively in the whole dataset and during the month of May 2013. The distributions are similar, while the mean and the standard deviation are just scaled by a factor around ten. For the next experiments, we decided to focus only on the month of May, as it is quite representative of the entire dataset.

So, after reducing our dataset, we have performed more detailed analysis to understand the current usage of taxis in Porto. Figure 3 represents the number of trips per day during the second week of May. For each day, we plotted in Fig. 4 the number

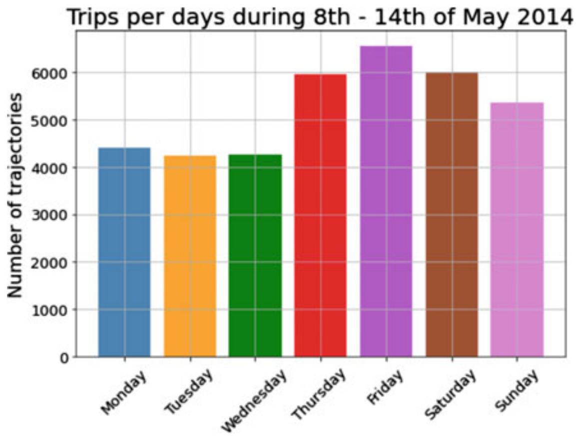


Fig. 3 Number of trips per day

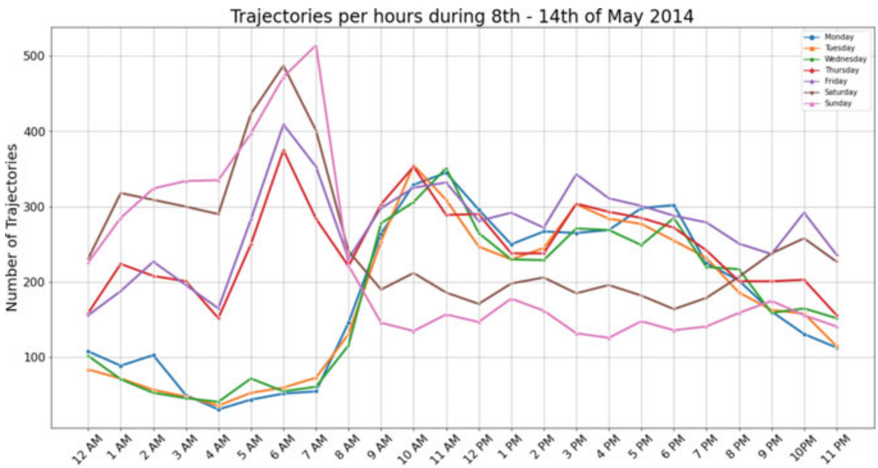


Fig. 4 Number of starting trips per hour

of starting trajectories per hour. Through these plots, we can already have different insights:

- It is noteworthy that days in which more taxi trips are done are from Thursday to Saturday. More in general, people prefer taking advantage of taxis during the long weekend, while less trips are done in the first part of the week.
- We can then divide the week into three chunks: from Monday to Wednesday, Thursday to Friday and the short weekend:

  – In the first group, we have few trips until 8 am, then a peak between 9 am and 11 am, and later another smoother peak between 5 pm and 7 pm. After that, the

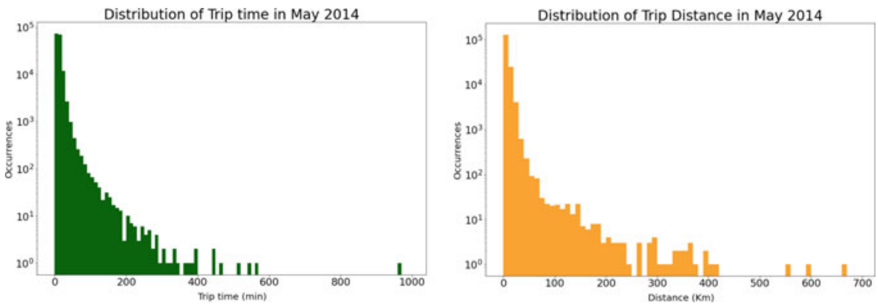


Fig. 5 Distribution of trips distance and duration

number of trips starts decreasing. Obviously, this trend is strongly connected with the working hours, as most people start working in that time slot, but the return to home is more diluted over time.

– In the second group, we can see different peaks: two of them are in the same correspondence of the ones before due to the round trip from home to work. On Friday, we can suppose that people come back home a bit earlier, since the peak is a bit shifted, around 3pm. There are also other two peaks during night, so it is conceivable that people usually go out in these two days, and need a taxi during the night.
– In the weekend, this behavior is more emphasized as the majority of taxi trips are done also during night. However, during the daylight we can register only few trips, and we noticed the absence of the two classical peaks, since most of people do not work.

We have gone further in the feature engineering process, and we have also extracted from our raw data the total distance of each trip and its duration. For this purpose, we have computed for each trip the distance between each pair of coordinates and then sum all the measurements. We have used the geodesic distance, which is more adapted than the Euclidean distance. The computation of the duration of a trip is mainly based on the number of [longitude, latitude] pairs as we have a record each 15 s.

The distribution of the duration and the distance of trips is given by Fig. 5.

As expected, we can notice that the two distributions are very correlated. The mean trip time is around 12 min while the mean trip distance is 6.5 km. The majority of the trips are quite short, and the third quartile of the distribution of the trip time is 15 min. In fact, a lot of trips take place inside Porto, and just few trips have a far destination from the starting point. Globally, the mean trip speed is around 27 km/h.

Finally, we presented in Fig. 6 a heatmap, which allows to understand which are the more common places in which people usually get a taxi, and which are their more common destination.

In Fig. 6, for the departures, we have highlighted, on the map, the first three hotspots. The first one is the Porto Campanhã station, that is the main rail station of

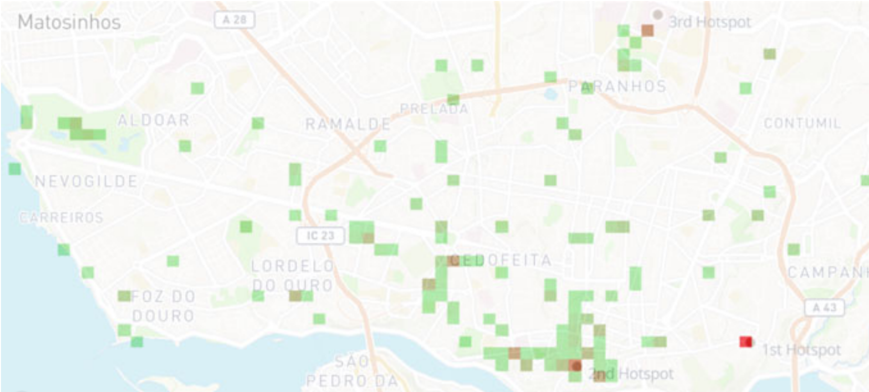


Fig. 6 Taxi trips departures in May 2013

Fig. 7 Taxi trips destinations in May 2013

the city, while the second one is the other, more ancient but more central, railway station, São Bento. The third hotspot is the University Hospital Center of São João, which is both the main hospital of the city and a medical school.

Figure 7 shows that the first two more requested destinations are the Porto Campanhã station and the Porto airport. The third popular destination is the other railway station, but more in general all the historic center is a common arrival point. We can notice that between points of departure and destinations there is a substantial difference: the former are scattered all over the city and they are placed both inside the center of Porto, but also outside “Via de Cintura Interna”, the ring road: the orange one on the Fig. 6, which encloses the city center. On the other hand, destinations are almost all concentrated in the city center, inside the ring road, and this is quite

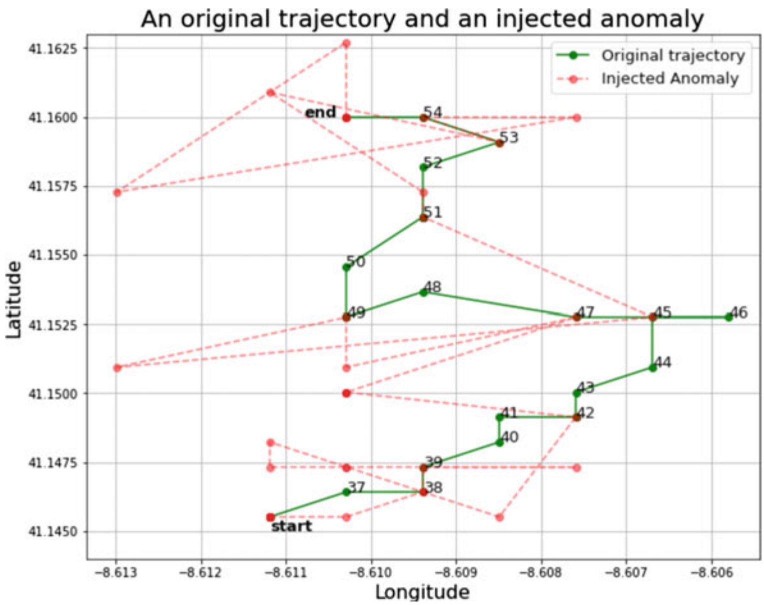


Fig. 8 Visualization of an injected outlier

reasonable. The few exceptions are the airport, the São João hospital and the Parque de Cidade at the west of the city, which is in Fig. 7, it located 4 cells.

3.2 Injected Trips Anomalies

As in [3], we generated several anomalies for evaluation since the dataset is unlabeled. The number of injected anomalies is around 5% of the size of the entire dataset. Figure 8 shows an example of injected anomalous trajectory having the same starting and ending point, as a real normal trajectory. Our objective is to compare two methods for the detection of these anomalous trajectories: Gaussian Mixture Variational Sequence Auto-Encode (GM-VSAE) and Anomalous Trajectory Detection using Recurrent Neural Network (ATD-RNN).

4 Experiments: GM-VSAE Versus ATD-RNN

The Porto Taxi Service Trajectory dataset is based on the GPS coordinates. In previous studies on GPS coordinates-based datasets, deep learning models are based on sequential learning such as: Long Short-Term Memory (LSTM) which is an RNN architecture that shows efficient results. LSTM-based approach has been used in recent studies, because of its ability of sequential learning, and it is also efficient in online detection of outliers. We have chosen to compare two sequential-based models: GM-VSAE proposed by Liu et al. [3] and ATD-RNN proposed by Song et al. [4]. Both models gave promising results on the Porto Service Trajectory dataset. The code of both models is available on github. We have run these models with the same pre-processing steps explained in [3, 4].

4.1 GM-VSAE

Trajectory data has become more complex due to the advancement in location-based technology. Traditional anomalies detection method cannot deal with the large and complex trajectory data. However, deep learning methods can provide solutions to this problem and are also able to detect anomalies in an efficient manner. Liu [3] characterized the anomalies as route switching and detour because the normal route definition cannot handle the variety, complexity and the sequential correlation between the routes traveled in the real-world scenarios. Trajectory is considered as detour anomaly if the route of the trajectory is longer than the normal route.

Anomalies detection in trajectory depends on the discovery of normal routes of trajectory. It is difficult to detect the normal routes because of the complexity of the transport system, and the variation of the route in different places. It is important to detect the anomalies online in the real-world scenarios, which is a challenging task due to the fast generation of the trajectory at massive scale. Detection of normal routes online is possible by assigning and updating the score according to the sequential information of the trajectory. However, it poses two problems: first is the complexity and variability of the trajectories data and second is the sequential correlation of the route traveled by the real-world trajectories. GM-VSAE provides the solution of these two problems and detects the trajectory anomalies online.

The architecture of the GM-VSAE consists of three components: route inference network, probability distribution of the routes and generative network. Route inference network converts the trajectory information into a vector in a latent space. RNN is used to capture the sequential information of the trajectory and handles the input in the form of the vector. Token embedded layer is introduced to convert the trajectory information into another vector. The sequential information is captured by RNN and represented in the latent space. In the second step, the probability distribution is used to measure the likelihood of the route to be considered as normal. It is a difficult task in the real-world scenario as the routes can be of different types, such as highway, street, ramps and others. To tackle this problem, C types of different routes are assumed for a given trajectory, here C is the hyper-parameter of the model. Two types of probability distribution (multinomial distribution and gaussian distribution) is used to discover the normal routes in the latent space. Multinomial distribution is used to model the probability of the type of the roads, while Gaussian distribution

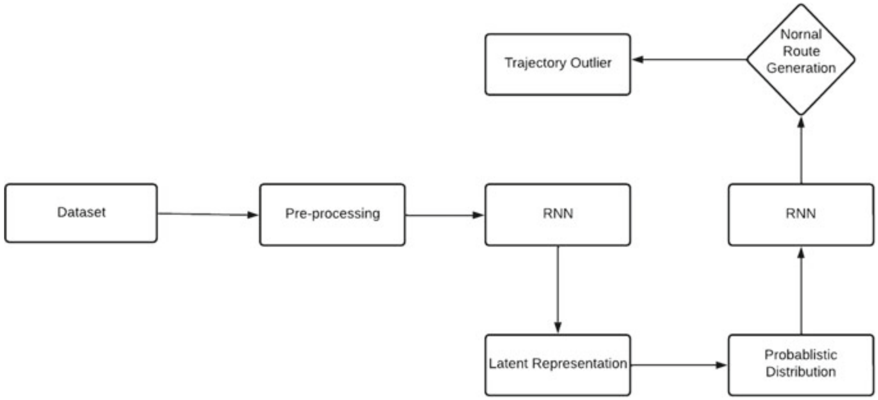


Fig. 9 Architecture of GM-VSAE

is used to measure the probability of the route type C traveled by the trajectories. Figure 9 shows the anomaly detection process using GM-VSAE.

In the next step, routes are generated from the probability distribution through RNN. RNN is used in generative network, because the trajectory data is sequential and each coordinate is linked with the previous coordinate in the trajectory data. The inputs are converted into the vectors, using the same embedded layer used in the route inference network before feeding it to the RNN. The routes are generated at each step of the trajectory data by getting the probability vector from multinomial distribution through a softmax function. The time cost of the GM-VSAE is proportional to C, i.e., the number of the component in gaussian distribution. Large number of gaussian components would slow down the online anomaly detection process. This problem can be solved by restricting the generation of the trajectory to one route, which has the highest probability in the trajectory data.

4.2 ATD-RNN

Anomalous trajectory detection can play an important role in many real-world applications such as: fraud detection, surveillance, etc. Most of the traditional methods do not consider the sequential information, because they are more focused on the historical information. Another disadvantage of the traditional methods is the data sparsity because they only take the given source and destination trajectories into consideration. To solve these problems, Song et al. [4] proposed ATD-RNN model, that detects the trajectory anomaly by trajectory embedding. The data sparsity issue is solved by considering more sources and destinations of the relevant trajectory.

The methodology of the ATD-RNN consists of three steps: data pre-processing, trajectory embedding and anomalous trajectory detection. In the data pre-processing stage, the trajectory data is converted into vectors fed as input to the embedding layer,

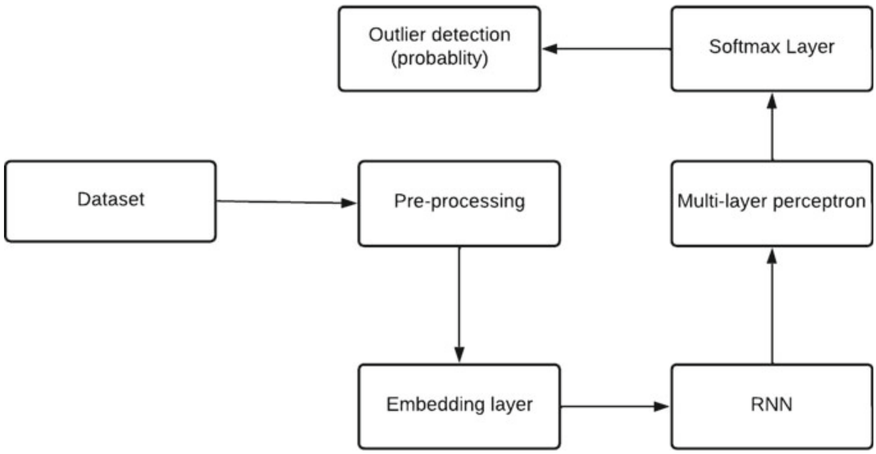


Fig. 10 Architecture of ATD-RNN

and after that to the RNN, which is used to capture the sequential information of the trajectory data. Softmax and multi-layer perceptron are used at the last step for the anomalous detection as shown in Fig. 10.

The trajectory data consist of the continuous GPS coordinates. Due to the large size of the trajectory data, we can learn trajectory embedding for every point and it would be difficult to generalize new point from the trajectory data. So, the trajectory data are divided into equal sized grids according to the hyper-parameters n and m, which are adjusted so that the size of the grid is about 100 m. After that each grid is uniquely labeled with an index number. The padding operation is performed on the trajectory to align them after obtaining the mapped trajectory as the length of the mapped trajectory is not equal. The main problem is that the anomalous trajectories are not common in the historical trajectory data. Anomalous trajectories are generated by disturbing some existing trajectory to solve this problem.

In the trajectory embedding step, stacked RNN is used to learn the trajectory embedding to find out the sequential information between the trajectory data. The mapped trajectories are fed into the stacked RNN sequentially. RNN can learn the sequential information through time and also memorize the information using the non-linear function which captures the trajectory characteristics in high-dimensional space. Dropout techniques are used to avoid the over-fitting problem [14]. The output state of the RNN is merged to get the trajectory embedding at the end. Multilayer perceptron is used to reduce the dimensionality of the trajectory embedding. After that, the result is fed into the softmax layer to generate the anomalous probability of the trajectory [4].

Table 2 Comparison of performance evaluation between ATD-RNN and GM-VSAE

	GM-VSAE	ATD-RNN
Precision	0.95	0.83
Recall	0.94	1.0
F1 score	0.954	0.90

4.3 Obtained Results

We have made the evaluation by comparing the results of the two defined models: GM-VSAE and ATD-RNN. The results of comparison between the two models are given in the Table 2. We can observe that GM-VSAE shows better results and it is more efficient in the real-world scenarios, as it can be very useful for online anomaly detection.

The two considered models have different architectures: GM-VSAE model uses data generation technique, while ATD-RNN has solved the data sparsity issue by inserting anomalies trajectories randomly in the historical trajectory. GM-VSAE shows better results and it is more generalized, because it uses probability distribution to find out the specific routes traveled by the trajectories.

5 Conclusion

In this paper, we performed a detailed analysis on the taxi service trajectory dataset to understand the usage of the taxi in Porto. After that we compared two sequential learning models: GM-VSAE and ATD-RNN. These two models capture the sequential information of the trajectories and can efficiently detect the anomalies in trajectories. Experiments on the taxi service dataset show that both GM-VSAE and ATD-RNN models give excellent results. Moreover GM-VSAE outperforms slightly ATD-RNN in terms of precision and F1-score.

References

1. Zheng Y (2015) Trajectory data mining: an overview. ACM Trans Intell Syst Technol 6(3) (Article 29). https://doi.org/10.1145/2743025
2. Meng F, Yuan G, Lv S et al (2019) An overview on trajectory outlier detection. Artif Intell Rev 52:2437–2456. https://doi.org/10.1007/s10462-018-9619-1
3. Liu Y, Zhao K, Cong G, Bao Z (2020) Online anomalous trajectory detection with deep generative sequence modeling. In: IEE 36th international conference on data engineering (ICDE), pp 949–960. https://doi.org/10.1109/ICDE48307.2020.00087
4. Song L, Wang R, Xiao D, Han X, Cai Y, Shi C (2018) Anomalous trajectory detection using recurrent neural network. In: Gan G, Li B, Li X, Wang S (eds) Advanced data mining and

applications. ADMA 2018. Lecture notes in computer science, vol 11323. Springer, Cham. https://doi.org/10.1007/978-3-030-05090-0_23
5. Li Q, Zheng Y, Xie X, Chen Y, Liu W, Ma M (2008) Mining user similarity based on location history. In: Proceedings of the 16th annual ACM international conference on advances in geographic information systems. ACM, p 34
6. Douglas DH, Peucker TK (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its Caricature. In: Cartographica: the international journal for geographic information and geovisualization, vol 10, Issue 2. University of Toronto Press Inc. (UTPress), pp 112–122. https://doi.org/10.3138/fm57-6770-u75u-7727
7. Clarke FH (1976) Optimal solutions to differential inclusions. J Optim Theory Appl 19(3):469–478. https://doi.org/10.1007/BF00941488
8. Huttenlocher DP, Klanderman GA, Rucklidge WJ (1993) Comparing images using the Hausdorff distance. IEEE Trans Pattern Anal Mach Intell 15(9):850–863. https://doi.org/10.1109/34.232073
9. Robinson MT (1990) The temporal development of collision cascades in the binary-collision approximation. nuclear instruments and methods in physics research section B: beam interactions with materials and atoms 48, no 1, pp 408–413. https://doi.org/10.1016/0168-583X(90)90150-S
10. Sankoff D, Kruskal J (1983) Time warps, string edits, and macromolecules: the theory and practice of sequence comparison. Addison Wesley, MA https://doi.org/10.1137/1025045
11. Suzuki N, Hirasawa K, Tanaka K, Kobayashi Y, Sato Y, Fujino Y (2007) Learning motion patterns and anomaly detection by Human trajectory analysis. In: IEEE international conference on systems, man and cybernetics, pp 498–503. https://doi.org/10.1109/ICSMC.2007.4413596
12. Lee J-G, Han J, Li X (2008) Trajectory outlier detection: a partition-and-detect framework. In: IEEE 24th international conference on data engineering, pp 140–149. https://doi.org/10.1109/ICDE.2008.4497422
13. Liu Z, Pi D, Jiang J (2013) Density-based trajectory outlier detection algorithm. J Syst Eng Electron 24(2):335–340. https://doi.org/10.1109/JSEE.2013.00042
14. Gal Y, Ghahramani Z (2016) A theoretically grounded application of dropout in recurrent neural networks. In: Advances in neural information processing systems, pp 1019–1027

Concept of Electronic Ship Electronic Record Book System Based on ISO 21745

Seongmi Mun, Gilhwan Do, and Kwangil Lee

Abstract MARPOL adopted the relevant amendment to allow the record book available electronically from October 1, 2020, Accordingly, ISO developed standard as ISO 21745 for electronic record book. In this paper, the concept of the system proposes to develop an electronic record book system corresponding with ISO 21745 based on international standards for ship networks.

Keywords Electronic record book · ISO 21745 · Ship standard network · ELB · Green ship

1 Introduction

A key element of the International Convention for the Prevention of Pollution from Ships (MARPOL) regulations is the recording of discharges associated with the prevention of pollution from ships. Therefore, a number of MARPOL Annexes require the recording of particular discharges. Traditionally, the format of these record books has been provided in hard copy by the Administration. However, as companies and shipowners increasingly focus on ways to operate in an environmentally responsible manner and aim to reduce the heavy burden associated with paperwork through electronic means, the concept of operational logs in an electronic format has become a popular consideration. Recently, IMO adopted amendments to MARPOL Annexes I, II, V, VI, and the NOx Technical Code will enter into force, enabling the

S. Mun (✉)
Seanus Co, 21, Mandeok-daero, Buk-gu, Busan, Republic of Korea
e-mail: kkang313@gmail.com

G. Do
C&P Korea Co. Ltd, 71, Jeoryeong-ro, Youngdo-gu, Busan, Republic of Korea
e-mail: ghdo@cp-korea.com

K. Lee
Korea Maritime and Ocean University, 727, Taejong-ro, Yeongdo-gu, Busan, Republic of Korea
e-mail: leeki@kmou.ac.kr

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_44

use of electronic record books (ERBs) in lieu of paper record books from 1 October 2020 [1].

Accordingly, the International Standard Organization (ISO)/TC (Technical Committee) 8/SC (Sub Committee) 11 began developing standards for electronic record systems in 2017 and established ISO 21745 as new standard with minimum technical specifications and operating requirements for ship electronic record books (ELRB) in 2019 [2–5]. Accordingly, in this paper, we propose a concept of an electronic ship electronic record book system corresponding to ISO 21745 based on the ship standard network.

2 Electronic Record Book

2.1 Definition and Necessity

Electronic record book means a device or system, approved by the Administration, used to electronically record the required entries for discharges, transfers and other operations as required under this Annex in lieu of a hard copy record book.

Currently, crew members on duty on board have to fill out the records four times a day at a fixed time, so the workload is considerable, and the contents cannot be identified according to the author's handwriting, which often hinders the continuity of work.

In addition, since the written record books are obligated to be stored on board for at least two years, the amount is very large, and it is very difficult to use the record book, so there is a demand for improving work efficiency and reducing resource waste by automating it.

The traditional record books are handwritten by visually checking the facility/ facility/sensor measurement data mounted on the ship through an indicator, and it will inevitably increase the time and cost as it is manually carried out throughout such as reporting and approval to superiors.

The record books must be kept for three years in accordance with IMO Convention and Article 44 of the Enforcement Rule of the Seafarers Act in Korea, but they are inefficient to store and manage record books, and there are always concerns about loss, contamination, and damage depending on the working environment.

Human errors such as errors and omissions occur in the process of checking the author's data and writing handwriting, resulting in a problem that the reliability of the recording information is deteriorated.

Accordingly, it is necessary to develop a standard interface that can automatically collect record books data from ship's various devices based on the technical specifications presented in ISO 21745 and to develop an ELRB system that can efficiently record, store, and manage data.

2.2 Sort of Record Book and Considerations

The types of electronic record books according to ship loading requirements are as follows:

- **Oil Record Book, parts I and II** (MARPOL Annex I, regulations 17.1 and 36.1) parts I oil fuel tank ballasting and washing water discharge, oil residue collection and disposal, oil mixture discharge, and other reasons for exceptional oil discharge
- parts II oil loading and internal transfer, waste disposal, waste disposal of separated waste tank, waste disposal, waste disposal of all waste tanks
- **Cargo Record Book** (MARPOL Annex II, regulation 15.1) when hazardous liquid substances or mixtures are accidentally discharged, the situation and reason of discharge, loading, and discharging operations
- **Garbage Record Book, parts I and II** (MARPOL Annex V, regulation 10.3) every discharge of garbage into the sea, every delivery of garbage to port waste reception facilities and every incineration operation with highlighting the position of the ship, the date and time of the operation, an estimate of the amount, and a description of the type of garbage
- **Ozone-depleting Substances Record Book** (MARPOL Annex VI, regulation 12.6); list of equipment containing ozone-depleting substances, amount of ozone-depleting substances filled in the facility, repair or management of the facility, atmospheric emissions of ozone-depleting substances, intentional or unintentional emissions, etc.
- **recording of the tier and on/off status of marine diesel engines** (MARPOL Annex VI, regulation 13.5.3)
- **Record of Fuel Oil Changeover** (MARPOL Annex VI, regulation 14.6)
- **Record Book of Engine Parameters** (NOX Technical Code, paragraph 6.2.2.7).

In this study, the Oil Record Book (ORB), parts I and II, and Garbage Record Book (GRB) are considered for the minimum software configuration. Defects caused by errors or deficiencies in the ORB are the third highest among all defects and are the main cause of massive financial and time damages. In addition, marine pollution caused by heavy oil, diesel, ship bottom wastewater, and other oil accounts for a large proportion of the total marine pollution accidents, and even a small amount of outflows are causing enormous damage. Over the past five years, the problem of marine pollution has become serious due to the outflow of pollutants caused by marine accidents, and record book helps preventing this.

Finally, I consider the Ballast Water Record Book (BRB). This is not a recommendation in MARPOL, but overseas competitors are using BWB as a management target. In addition, considering the ecosystem disturbance caused by ballast water, systematic management of ballast water and Ballast Water Treatment System (BWTS) is necessary. This is because it is estimated that 10 billion tons of ballast water travel to each country every year, and more than 7,000 species of marine life travel.

3 Proposed Electronic Record Book System

3.1 Concept of the System

This study aims to develop and demonstrate the Korean first ELRB system that collects and monitors data to meet international standard requirements for ship networks and automatically records and manages record books for ships based on the latest international standard ISO 21745. First of all, the schematic concept of ELRB proposed for this purpose in this paper is as follows (Fig. 1).

The kind of data to be collected needs to be defined in accordance with the traditional record book and ISO 21745 standards and user requirements. Accordingly, it is necessary to classify automatically collected data and data that must be manually entered. This system configuration is constructed based on data that is automatically collected and stored. The data collection system (DCS) collects data from the engine sensor and the navigation sensor through the Machine/Navigation Sensor Interoperability Gateway (MNIG). The MNIG receives data from navigation sensors such as GNSS, Gyro Class, AIS, and VDR based on IEC61162-1/2 and machine sensors such as main engine, generator, and boiler based on IEC 61162–450 standards. The MNIG can communicate with the data collection system through the IEC 61162–450 communication module to store and monitor the converted data. The database will be designed in accordance with ISO 847, ensuring that data collected from MNIG can be shared in a safe and efficient manner. The ELRB main server communicates with the DCS using the IEC 61162–450 communication module to transmit the data required for system configuration. ELRB will provide a Web-based service, so it is necessary to configure a Web server.

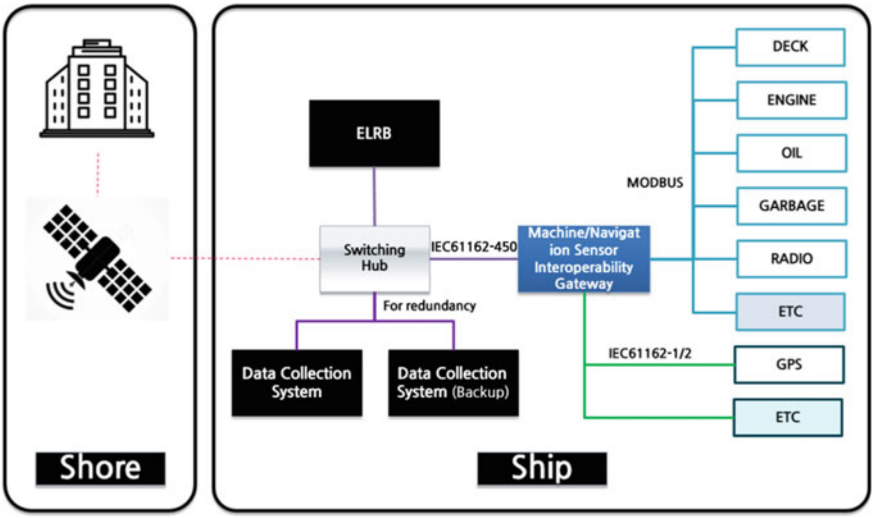


Fig. 1 System configuration

3.2 Analysis of Requirement of ISO 21745

System requirements are divided into four categories: general requirements, functional requirements, human–machine interface, and system updates.

Since the purpose of this study is not to develop a system, but to design software to develop, only the contents that directly affect the software configuration among the four requirements are considered.

General requirements are related to power supply, and human–machine interface is excluded because they are requirements from an ergonomic perspective. System availability is also excluded because it is a requirement related to hardware configuration. Functional requirements include data store, record management, system output, validation, and system availability, and system updates are requirements related to updating the corresponding electronic records. After analyzing ISO 21745, the minimum requirements for designing a software prototype were identified as follows (Table 1).

In future, we will further investigate the standard documents and user requirements needed to configure ELRB to conduct detailed software design.

Table 1 Minimum requirements for software

Category	Sub category	Item
Functional requirements	Data storage	• Whether ELRB and traditional record book information match • Whether UTC and latitude/longitude record and store • Whether SW internal clock is synchronized with a UTC source such as GPS • Whether ELRB is recorded and stored in English • Whether readable font • Whether ELRB data is classified into 1) automatic collection data, 2) register data, 3) signed register data, and 4) editing history data • If the ELRB has auto-recording, whether the data collected automatically • If not automatically recorded in the main storage, whether it is displayed on the screen • Whether data collected automatically is provided with manual input if it is not automatically recorded in the main store • If ELRB cannot record data, whether the 'data shall be recorded in an official paper logbook' sign is permanently displayed

(continued)

Table 1 (continued)

Category	Sub category	Item
	Record management	• Whether only authorized persons on board the vessel can complete the ELRB entry • Whether automatically collected data and editing history data cannot be modified • Whether only authorized persons can edit or modify register data • Whether signed register data can only be modified or modified by the master • Whether accepting the record book data as signed record book data is only permitted for a master with full access to view, modify, and sign ELRB data • Whether editor, date, and time of modification are indicated for all modifications • Whether reasons for making changes for master verification and who records the changes • If changes to each section are required, the original and all amendments to each section must be maintained and visible
	System output	• Whether the output data is represented in a file format that prevents it from being modified or collected • Whether the document is provided as a PDF
	Validation	• Whether the name of the person authorized to the ELRB or the identification of the other official person who performed the recording activity is mentioned • When (UTC, date, and time) each record is described and by which authorized person • Whether the content of the audit logging is (1) creating data items, (2) editing or modifying data items, and (3) deleting or verifying data items • Whether audit logs are accessible and exported • Whether log entries can be filtered by 'activities executed by authorized person or in a specific time window'
System updates	System updates	• Whether it provides a means of displaying the current software version • Whether means are provided for replacing software on board systems or installing updates

4 Conclusion

In this paper, the concept of the entire system is proposed to develop an ELRB corresponding to ISO 21745 based on international standards for ship networks. In future, we will conduct a detailed design of software by further investigating the standard documents and user requirements required to configure ELRB.

Acknowledgements This work is supported by Korea Institute of Marine Science and Technology Promotion(KIMST) funded by the Ministry of Ocean and Fisheries, Korea(20220531). Also this work was supported by the National IT Industry Promotion Agency (NIPA), grant funded by the

Korean government Ministry of Science and ICT (MSIT). Grant No. S1712-22-1001, for ISO21745 based Ship Electronic Record Logbook System.

References

1. Standard Club Homepage. https://www.standard-club.com/knowledge-news/news-guidelines-for-the-use-of-electronic-record-books-erbs-under-marpol-1490/. Last accessed 20 Oct 2022
2. Resolution MEPC.312(74) (adopted on 17 May 2019) Guidelines for the use of electronic record books under marpol
3. Resolution MEPC.314(74) (adopted on 17 May 2019) Amendments to the annex of the international convention for the prevention of pollution from ships, 1973, as modified by the protocol of 1978 relating thereto
4. Resolution MEPC.316(74) (adopted on 17 May 2019) Amendments to the annex of the protocol of 1997 to amend the international convention for the prevention of pollution from ships, 1973, as modified by the protocol of 1978 relating thereto
5. Resolution MEPC.317(74) (adopted on 17 May 2019) Amendments to the Nox technical code 2008

The Use of Latent Semantic Analysis for Political Communication: Topics Extraction for Election Campaigns

Grassia Maria Gabriella, Marino Marina, Mazza Rocco, and Stavolo Agostino

Abstract In the era of the digital revolution, the ability to share and analyze user-generated content has opened new challenges for researchers. Especially in the political field, probing public opinion and understanding how online users express themselves on a given political issue is becoming increasingly central to parties and politicians. The contribute shows a strategy for extracting the major issues of discussion on which to base the campaign of politicians running for election. To do this, latent semantic analysis was applied to contents produced within a local Facebook group. At the end of the work, the themes that will be the basis for the agenda setting of the political class are displayed.

Keywords Latent semantic analysis · Political communication · Topic-based approach

1 Introduction

Micro-blogging platforms and social networks play a predominant role to detect opinions and attitudes on relevant topics [12]. The prospects for the use of social media appear to be promising in the political context because of the possibility of fostering public participation and democracy [3]. Therefore, social media allow increasing the

G. M. Gabriella (✉) · M. Marina · S. Agostino
University of Naples "Federico II", Naples, Italy
e-mail: mariagabriella.grassia@unina.it

M. Marina
e-mail: marina.marino@unina.it

S. Agostino
e-mail: agostino.stavolo@unina.it

M. Rocco
University of Campania "Luigi Vanvitelli", Caserta, Italy
e-mail: rocco.mazza@unina.it

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_45

political participation of citizens and political institutions [19]. Political institutions use social media to derive information from citizens on certain policy issues or to probe public opinion, thus establishing direct contact with the public [23]. The use of platforms allows candidates to mobilize voters and build communities [20], especially during election campaigns to discuss political issues, promote certain issues, and accentuate salient features of candidates [14]. On the other hand, citizens discuss and debate political issues, consult profiles of politicians or parties, and create a direct communication channel with candidates. It is becoming increasingly important to collect, monitor, and analyze user-generated political information on social media. On the issue, the study aims to analyze and explore the content produced by citizens in the Facebook group called “Open Succivo” to identify the main topics of discussion, according to latent semantic analysis, on which to base the campaign for elections. The rest of the paper is organized as follows. In Sect. 2, we present a brief review of research in this field. In Sect. 3, we describe our proposal to identify the main topics with LSA. In Sect. 4, we show the results obtained by analyzing a dataset of contents published on Facebook. At least, in Sect. 5, we define the conclusions with some observations and future research orientations.

2 The Theoretical Framework

Latent semantic analysis (LSA) is a factorial technique to represent the meaning of terms defined in the context of use in a collection of documents, through a matrix of reduced dimensionality [4]. LSA has applications in the fields of information retrieval, artificial intelligence, psychology, cognitive science, education, and text mining [6]. Specifically, to the research carried out, it is also applied in the political communication field to extract the main topics of discussion. Particularly [22] adopted the LSA on the transcripts of the 2016 presidential debates between Hillary Clinton and Donald Trump to identify political issues. Hacker et al. [13] investigate the war and peace speeches of two Iranian leaders following changes in Iranian government communication. Finally, Conover et al. [2] carried out a study to verify whether political candidates define the relevant topics on which to base the election campaign.

The LSA technique represents data through a vector model space and projects a matrix of terms-documents in a space of factors with reduced dimensionality, as well as identifying the relationship between its component terms [21]. The starting point is the *terms x documents* matrix X, where the rows are associated with words, while on the columns, there are the documents or generally text segments. In the matrix, each cell contains the frequency with which the words appear in the documents denoted by the columns (*term frequency*). Usually, this kind of matrix is too sparse, so it is necessary to transform the distribution of terms according to the weight. Thus, the frequency is weighted by a function that expresses the importance of the word in documents [15]. Then, LSA applies singular value decomposition (SVD) to the matrix X:

$$X = U \sum V^T \tag{1}$$

where U is an orthogonal matrix of term eigenvectors, V is an orthogonal matrix of documents eigenvectors, and $\sum$ is a diagonal matrix of singular values where the remaining cells are zeros [11].

A low-rank approximation matrix will replace the original matrix based on the SVD of X [9], after selecting the so-called rank, or the number of dimensions to define the reduced matrix. Therefore, the SVD reproduces the matrix X using a space of latent semantic dimensions. These dimensions explain variability in term-document occurrences, and they are quantified in singular values of the diagonal matrix [6]. So, selecting the k number of dimensions according to the highest value of singular values, the matrix X_k is truncated:

$$X_k = U_k \sum k V_k^T \tag{2}$$

The matrix X_k is a least-squares best approximation of the original matrix, minimizing the sum of the squared differences between the elements of X and X_k [7]. The matrix X_k is created by setting equal to zero all elements except the first k or columns of term vectors in U, the first k singular values in $\sum$, and the first k elements or columns of document vectors in V. The columns of U and V are orthogonal, but the rows are not orthogonal [16]. According to the orthogonal characteristic of factors, words have high relations with terms that are in the factor but have little relation with words in others [17]. The k-dimensional vector space is the base for the semantic structure used by the LSA. In general, types similar in meaning are "near" each other in the space even if they never co-occur in a document, and documents similar in conceptual meaning are near each other even if they share no types in common [1]. The technique can be represented in a geometrically in Fig. 1 [5].

From Fig. 1, the terms and the documents are visualized like vectors in the space to k-dimensions; moreover, the axes produced from the SVD are a linear combination of the terms. As you can see from the representation, the dimensions derived from the SVD are orthogonal to each other, but the terms are not. This results in term vectors not being independent of each other but the position they occupy reflects correlations in their use in documents.

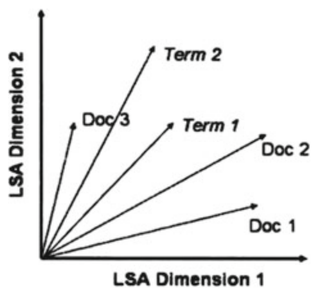


Fig. 1 Geometric representation of latent semantic analysis Dumais (*Source* Dumais [5], p 194)

In this regard, vectors are compared with each other using the cosine similarity measure. The cosine measure is used to understand which vectors terms and document vectors are most similar to each other and established a specific threshold. It is given by the point product of the vectors relative to the product of their size [10] and is defined by the following formula:

$$\cos(\vec{u}, \vec{v}) = \frac{\vec{u} \cdot \vec{v}}{||\vec{u}||_2 \cdot ||\vec{v}||_2} = \frac{\sqrt{\sum_i (\vec{u} \cdot \vec{v})}}{\sqrt{\sum_i (\vec{u}^2)} \cdot \sqrt{\sum_i (\vec{v}^2)}} \tag{3}$$

Cosine similarity between documents shows the angle of two document vectors in document space. Cosine value must be ranging from −1 to 1: when the cosine measure is equal to 1, the angle between the vectors indicates that the closer they approach the value 1 the more they will be similar; and if the value approaches 0, the terms between the two vectors have no similarity and may also mean that they are unrelated. Since the cosine measure does not consider vector lengths, it is possible to compare terms or documents of different lengths [10].

3 Methodology

For investigating what are the relevant issues on which to base the election campaign for the local elections in the city of Succivo, we analyze a Facebook group in which the town's users most discuss issues related to the proper functioning of the city, presenting solutions and alternatives to solve them, or highlighting critical issues. We choose Facebook for our analysis because on the platform, there are numerous contents produced by the users, and the data is semi-public in nature, meaning that they can be produced and collected from public profiles or groups. More specifically, in the political context, it is useful to define a set of politically relevant Facebook groups and pages in order to create a database produced by users, where posts are not limited in their length. To understand the above, we selected the Facebook group called "Open Succivo", a public group that contains 7144 members, collecting posts and comments published by citizens. We extracted data from January 2019 to December 2021—three years—using the scraping software CrowdTangle. Excluded from the selection of units are videos, images, and links to articles and/or other external groups. The next step involving the extraction of elements was done by creating a matrix containing the following variables: date of publication, gender of the subjects, type of content, and textual element. Through the software, we extracted the usernames of the users within the group, so we could define by name the gender of the person participating in the discussion: male (60%) and woman (40%). In the end, we had 3599 contents, composed of 839 posts and 2760 comments.

Having constructed the starting matrix, it is necessary to go through all the cleaning and pre-processing of the text. The goal is to transform the unstructured data (i.e. the text element) into structured data, which will be subjected to statistical-mathematical operations. There are different steps:

Tokenization and parsing: a document can be seen as a sequence of characters, resulting in a set of distinct strings (tokens) separated by spaces or punctuation marks. So, the text is decomposed into its constituent components according to a particular encoding called bag-of-words (BoW) that represents documents as arrays containing the occurrences of individual words.

Normalization: after segmenting the text into tokens, it is reprocessed through a series of techniques to reduce language variability, thereby improving the effectiveness of subsequent analytical steps. It is often necessary to change the characters of all terms to lowercase, going on to define *normalization.*

Elimination of stop-words: these words do not contribute to the meaning of the documents, so they do not have a significant value. Through the procedure of eliminating stop-words, it is possible to remove words that are useful in composing a meaningful sentence but that are isolated from the context (e.g. prepositions and articles) and special characters (e.g. hashtags and emoticons).

Grammar tagging: the process of branding a word in a corpus as corresponding to a particular part of the discourse. This stage turns out to be central because it allows for the recognition of part of speech (POS) functional to the identification of word categories. We selected the nouns, the verbs, and the adjectives.

Lemmatization: this phase involves tracing each inflected form back to the lemma. In textual analysis, it indicates that we consider the infinitive for verb forms, the singular for nouns, and the masculine singular for adjectives.

At the end of the pre-treatment phases, we have 10,258 tokens, 4720 types, and 3599 documents. We created the *terms x documents* matrix, where the dimension is 4720×3599. As we say, this matrix is too sparse; so, we create a weighted matrix according to the *tf-idf*, the term frequency—inverse document frequency [10]. Here, we applied the latent semantic analysis for extracting the topics.

4 LSA Outcomes

The following topics by LSA are identified as significant. We reported for each theme the geometric representation in two dimensions (Fig. 2).

The first topic identified pertains to the management of the local government and administration. The discussion involves not only the elected officials, including the current mayor ("Papa") and ex-mayor ("Colella"), but also the citizens' ongoing critique of the administration's handling of tenders. These critiques highlight a lack of transparency in the bidding process, resulting in the predictable victory of local businesses. Over the three-year period studied, the tenders were released to address various issues in the region, such as improving road safety and building sidewalks. The residents of Succivo are demanding that the local government approve projects and make changes to improve the city's liveability, as it is currently experiencing a state of disruption ("dissesto") (Fig. 3).

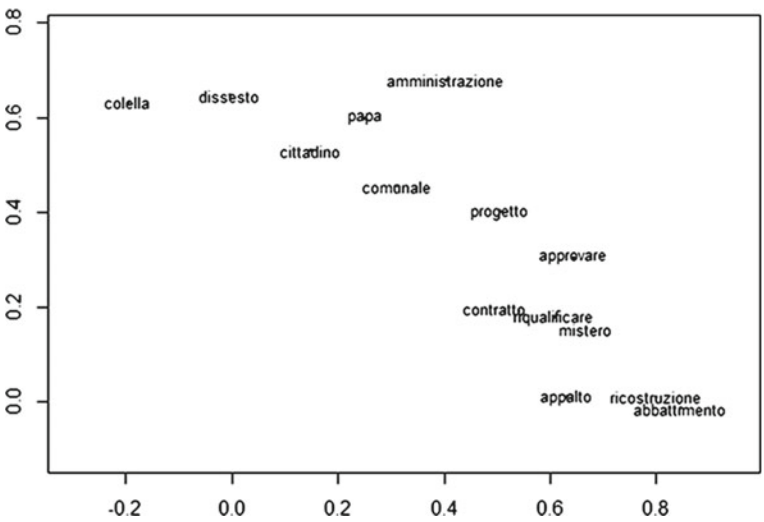


Fig. 2 Terms associated with the topic administration ("amministrazione")

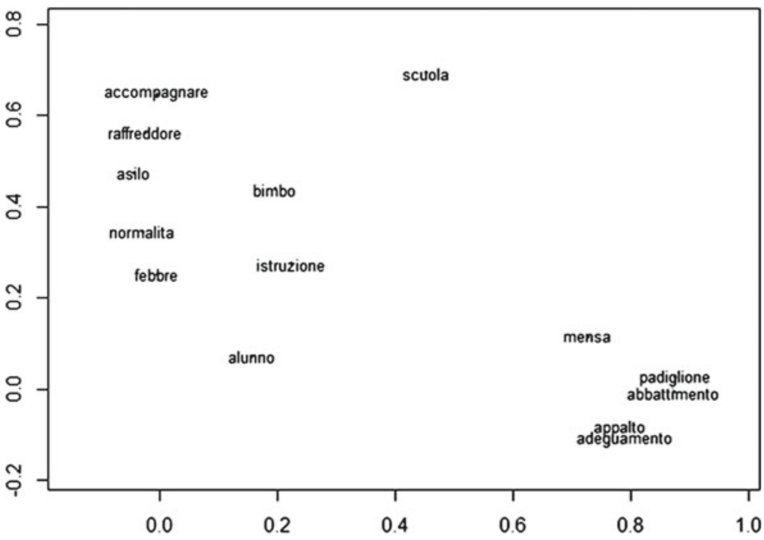


Fig. 3 Terms associated with the topic school ("scuola")

The topic of school ("scuola") plays a central role in the discussion on the Facebook group. Users complain that parents also accompany their children with symptoms such as cold ("raffreddore") and fever ("febbre") especially during the COVID-19 period. The term normalcy ("normalità") refers to the need to return to a school in attendance, as before the pandemic emergency. Also, the terms canteen ("mensa"),

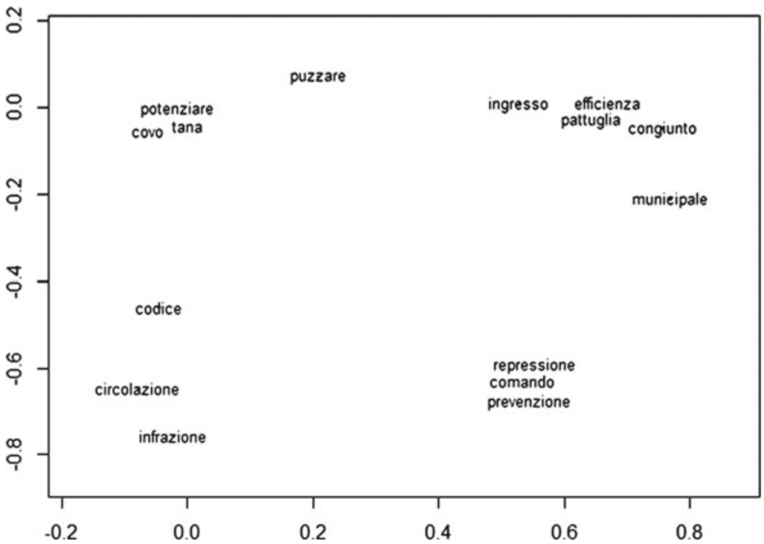


Fig. 4 Terms associated with the topic security ("sicurezza")

pavilion ("padiglione"), and abatement ("abbattimento") refer to a particular incident during which one of the roofs of a local elementary school collapsed. The related timing of securing it, as well as the fear of having places that were unsuitable for children's education, were the subject of criticism by users of the group. The latter issue has been of strong discussion as according to citizens, and the administration has been unable to manage the school safety work, making various mistakes (Fig. 4).

The issue of security ("sicurezza") is central to the public debate. Residents of the city complain of poor security and poor control in the area; in fact, they demand that the mayor implement control policies to prevent dangers in the area. Specifically, the terms refer to the COVID-19 period during which it was necessary to maintain rules such as distancing, wearing individual devices, and the possibility of visiting only relatives ("congiunto") during the first phase of the pandemic. Respect for the rules and the issue of control are central to the public debate: Succivo residents complain about the lack of compliance with rules and a lack of criminal sanctions ("infrazione") and controls by law enforcement ("potenziare") agencies (Fig. 5).

The last topic excerpted must do precisely with the environment ("ambiente"). The city of Succivo is one of the 90 municipalities belonging to the territory of the so-called "Terra dei fuochi". It refers to that territory ("territorio"), between the province of Naples and the south-western area of the province of Caserta, affected by the phenomenon of illegal landfills and/or the uncontrolled abandonment of urban waste and special ones, often associated with the combustion of the same. Succivo falls within a critical area for the presence of toxic ("tossico") fires and illegal waste disposal. Citizens complain of poor management of the situation, along with other municipalities ("Frattamaggiore," "Casoria", "Gricignano") and an administration

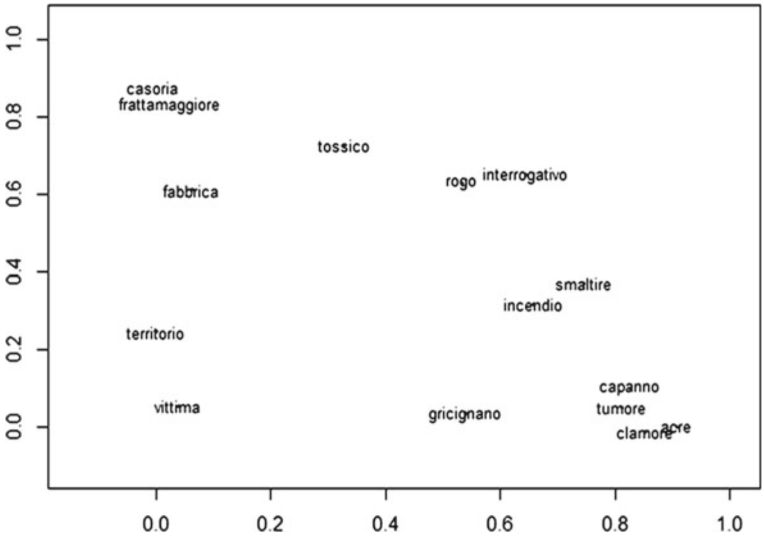


Fig. 5 Terms associated with the topic environment ("ambiente")

that is absent in addressing environmental issues. In addition, the members of the group constantly report the numerous victims ("vittime") of cancer ("cancro") death.

5 Conclusions and Future Remarks

At the end of the work, we can check what were the topics of greatest discussion within the "Open Succivo" Facebook group. This approach allows political candidates to create an election campaign based on the critical issues and suggestions expressed by the citizens. For political and party representatives, it is important to identify and monitor political arguments, in order to understand what issues are expressed by users on which to base the election campaign, creating policies that are in line with the demands and issues expressed by the citizens. By doing so, politicians can understand the public opinion as it positions itself on a given issue, as well as analyzing the consequences expressed by users [18]. The study fits into the theoretical field through which local political communication is treated with a statistical technique such as LSA, which allows the identification of discussion topics through an association between terms. The paper aims to represent a new application example of how a statistical technique such as LSA can be worked on the analysis of numerous short texts (posts and comments) extracted from social media. In fact, LSA requires a large number of texts to perform SVD analysis. By leveraging the capability to process large amounts of data, including the content generated by internet users, it becomes feasible to delineate the semantic domain more accurately, enabling a greater variety of contexts in which words can co-occur with one another

[8]. The choice to use LSA as an analysis technique depended on the fact that it does not require much statistical background [17]: the language used to explain LSA is similar to the general linear model. In addition, LSA requires relatively less computing power than other methods because most estimates are computed on eigenvector matrices [22]. The latent semantic analysis has some restraints. First, it is a matrix size reduction technique; therefore, it is not based on the study of probability distributions (such as Latent Dirichlet Allocation). Moreover, the process of identifying the number of factors is not statistically determined but is the result of the researcher's reasoning. Finally, polysemy is partially treated in LSA due to the characteristic of orthogonality. To overcome the limitations described above, such as the simple matrix reduction and the problem of polysemy, several technical updates have been proposed, such as probabilistic latent semantic analysis (pLSA) [17].

References

1. Berry MW, Young PG (1995) Using latent semantic indexing for multilanguage information retrieval. Comput Humanit 29(6):413–429
2. Conover MD, Gonçalves B, Ratkiewicz J, Flammini A, Menczer F (2011) Predicting the political alignment of twitter users. In: 2011 IEEE third international conference on privacy, security, risk, and trust and in 2011 IEEE third international conference on social computing, pp 192–199
3. Creighton JL (2005) The public participation handbook: making better decisions through citizen involvement. Wiley
4. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R (1990) Indexing by latent semantic analysis. J Am Soc Inf Sci 41(6):391–407
5. Dumais ST (2005) Latent semantic analysis. Ann Rev Inf Sci Technol 38:188–230
6. Evangelopoulos N, Zhang X, Prybutok VR (2012) Latent semantic analysis: five methodological recommendations. Eur J Inf Syst 21(1):70–86
7. Evangelopoulos NE (2013) Latent semantic analysis. Wiley Interdisciplinary Rev: Cognitive Sci 4(6):683–692
8. Foltz PW (1996) Latent semantic analysis for text-based research. Behav Res Methods Instrum Comput 28(2):197–202
9. Gansterer WN, Janecek AGK, Neumayer R (2007) Spam filtering based on latent semantic indexing. Tech Rep
10. Gefen D, Endicott JE, Fresneda JE, Miller J, Larsen KR (2017) A guide to text analysis with latent semantic analysis in R with annotated code: studying online reviews and the stack exchange community. Commun Assoc Inf Syst 41(1):21
11. Golub GH, Van Loan CF (2013) Matrix computations. JHU Press
12. Grassia MG, Marino M, Mazza R, Stavolo A (2022) Analysis of the public debate on DDLZan on Twitter: an application of the structural topic model in 16th International conference on statistical analysis of textual data JADT
13. Hacker KL, Boje D, Nisbett VL, Abdelali A, Henry N (2013) Interpreting Iranian leaders' conflict framing by combining latent semantic analysis and pragmatist storytelling theory. In: Political communication division of the national communication association annual conference
14. Kobayashi T, Ichifuji Y (2015) Tweets that matter: evidence from a randomized field experiment in Japan. Polit Commun 32(4):574–593
15. Landauer TK, Foltz PW, Laham D (1998) An introduction to latent semantic analysis. Discourse Process 25(2–3):259–284

16. Landauer TK, McNamara DS, Dennis S, Kintsch W (2011) Handbook of latent semantic analysis. Routledge
17. Lee S, Song J, Kim Y (2010) An empirical comparison of four text mining methods. J Comput Inf Syst 51(1):1–10
18. Sobkowicz P, Kaschesky M, Bouchard G (2012) Opinion mining in social media: modeling, simulating, and forecasting political opinions in the web. Gov Inf Q 29(4):470–479
19. Stieglitz S, Dang-Xuan L (2013) Social media and political communication: a social media analytics framework. Soc Netw Anal Min 3(4):1277–1291
20. Stier S, Bleier A, Lietz H, Strohmaier M (2018) Election campaigning on social media: Politicians, audiences, and the mediation of political communication on Facebook and Twitter. Polit Commun 35(1):50–74
21. Underhill TN (2007) An introduction to information retrieval using singular value decomposition and principal component analysis
22. Valdez D, Pickett AC, Goodson P (2018) Topic modeling: latent semantic analysis for the social sciences. Soc Sci Q 99(5):1665–1679
23. Zeng D, Chen H, Lusch R, Li S (2010) Social media analytics and intelligence. IEEE Intell Syst 25(6):13–16

A Data Analytics Methodology for Benchmarking of Sentiment Scoring Algorithms in the Analysis of Customer Reviews

Tesneem Abou-Kassem, Fatima Hamad Obaid Alazeezi, and Gurdal Ertek

Abstract Due to the digitalization, there exists an increased amount of user-generated content on the Internet, where people express their opinions on various topics. Sentiment analysis is the statistical and analytical examination of human emotions and opinions regarding a certain subject. Our study extends the literature by developing a data analytics methodology for the benchmarking of sentiment scoring algorithms in the context of online customer reviews. We demonstrate the applicability of the methodology using Amazon product reviews as the source data. Analyzing text-based content such as Amazon customers' reviews through text analytics and sentiment analysis can help Amazon and other online retailers to discover valuable actionable insights regarding their products. The contributions of this study are twofolds: to examine the predictive power of machine learning (ML) algorithms with respect to predicting sentiment scores and to analyze patterns in the differences between scores obtained from different sentiment scoring algorithms.

Keywords Online customer reviews · Sentiment analysis · Text analytics · Machine learning · Gap analysis

T. Abou-Kassem · F. H. O. Alazeezi · G. Ertek (✉)
United Arab Emirates University, Al Ain, UAE
e-mail: gurdal@uaeu.ac.ae

T. Abou-Kassem
e-mail: 201640019@uaeu.ac.ae

F. H. O. Alazeezi
e-mail: 201601253@uaeu.ac.ae

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_46

1 Introduction

Massive amounts of digital data and information are captured almost every moment, regarding almost every aspect of our lives. Human behavior is strongly influenced by sentiments/emotions and beliefs, which affect judgments and decisions. Considering how different people perceive and propagate the world and its various aspects can significantly influence our decisions [1]. In the context of e-commerce, which is the domain of interest in this paper, analyzing sentiment is very important to understand customers' needs and wants, as well as improving products or services delivered to customers. Forums, blogs, customer reviews, social networks, all coexist in the ever-growing social media world, all of which can be analyzed through techniques referred to as "Sentiment Analysis" [1]. Sentiment analysis is the process of determining whether a given text is positive, negative, or neutral. It can also be used to analyze a variety of different types of data, including social media posts, reviews, articles, and more [2]. There are a few different ways to perform sentiment analysis. One common way is to use a lexicon, which is a list of words that categorize text with common characteristics. Sentiment analysis determines whether the opinion is positive or negative for a topic or entity on the Internet, for topics such as economy and finances, and entities such as movies and products. The majority of social media data is unstructured because of the variety of available formats for messages, posts, and other content, and due to the easy accessibility of the social platforms. To make a decision, users typically search for and take as reference others' reviews, opinions, and experiences which can yield valuable information for users, but can also be used to mislead them. Motivated by the importance of online customer reviews and their impact on consumer decisions, our study contributes to the literature by introducing a data analytics methodology for the benchmarking of sentiment scoring algorithms, in the context of online customer reviews. The applicability of the introduced methodology is demonstrated through empirical analysis in a case study.

2 Literature Review

Our study is based on data obtained from Harvard Dataverse [3], which was originally collected by Chatterjee et al. [4]. We refer to this dataset as Dataset A. The authors carried out outlier detection and sentiment analysis using the data, as a case study of Amazon customer reviews. It shows a statistics-based outlier detection and correction method (SODCM) that finds reviews and fixes their star ratings. This makes sentiment analysis algorithms better without degrading the quality of the data. Fang and Zhan [5] discussed the process of sentiment polarity categorization using both sentence-level and review-level categorizations. Furthermore, they split their work into three phases; their main work was in phases 2 and 3, where they conducted the sentiment score. The authors then conducted tests to compare and evaluate the results of different algorithms for scoring sentiment. Naseem et al. [6] present a large-scale benchmark Twitter dataset for COVID-19 sentiment analysis. They evaluated and labeled the sentiment scores as positive, negative, and neutral using the TextBlob algorithms only. As part of their sentiment classification task, they used different machine learning methods and deep learning-based classifiers. Onan [7] presented a deep learning-based architecture for sentiment analysis using Twitter product reviews, which combined glove-weighted TF-IDF word embedding with a CNN-LSTM-based architecture. Also, the author discussed how words and sentences make sense based on how they are arranged in a dictionary. This is how the orientation of a text document is found. For machine learning-based classification models, the author used labeled datasets as training sets for supervised learners. Rezaeinia et al. [8] introduce Improved Word Vectors (IWVs) as a new technique to increase the accuracy of the pre-trained word embeddings in sentiment analysis. Part-of-Speech (POS) tagging techniques, lexicon-based approaches, word position algorithms, and Word2Vec/GloVe approaches were all used in their study. Mowlaei et al. [9] state that to get a better idea of how the public feels about a campaign, it is important to look at written reviews by extending two lexicon generation methods for aspect-based issues: one that uses statistical methods and the other that uses genetic algorithms to create the sentiment analysis. Al-Shabi [10] uses lexicon-based sentiment analysis as the primary method of analysis. The mentioned study focuses on VADER [11], SentiWordNet, SentiStrength, the Liu and Hu opinion lexicon, and AFINN-111, which are among five of the most important and well-known sentiment analysis lexicons/algorithms for Twitter data. The author's results show how well these lexicons/algorithms perform at classifying the polarity of tweets by comparing the overall accuracy of classification with the F1-measure.

3 Methodology

The objective of the study presented in this paper is to extend the methodological and practical body of knowledge in sentiment analysis, in the context of online customer reviews. To this end, an application-oriented data analytics methodology has been developed, documented, and implemented using real-world data.

The methodology developed for and applied during the study is provided in Fig. 1. Firstly, the source data are processed and prepared for analysis. This preparation step also includes the engineering, computation, transformation, and generation of existing and new attributes. Secondly, standard text analytics steps are followed to analyze the text corpus (collection), which, in our study, is the collection of online customer reviews at Amazon.com. Thirdly, word frequency tables are used in conjunction with sentiment scores for predictive analytics to benchmark the various algorithms. Finally, gaps between scores generated by two sentiment scoring algorithms are analyzed.

Using Natural Language Processing (NLP) as the text analytics technique, Python programming language, and KNIME data analytics platform to explore trends and

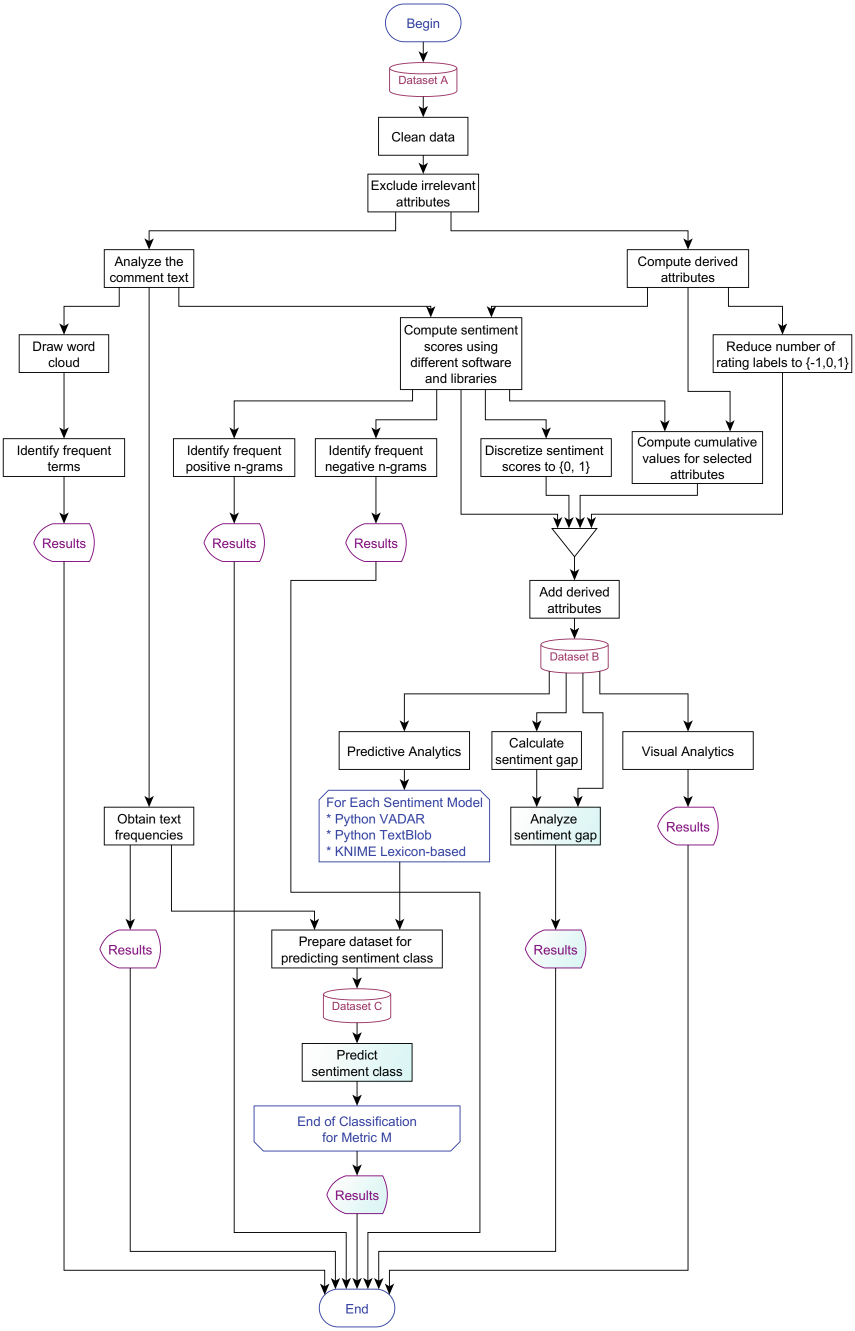


Fig. 1 Data analytics methodology applied in the presented research study

sentiment analysis in Amazon's customer reviews toward a specific product can enable various insights. The presented techniques/tools can help in understanding what/how customers feel about their purchases, what they like and dislike about products, which factors are highly associated with sentiments, which sentiment scoring algorithms generate scores that have the highest performance with respect to predictability, and how gaps between gaps between sentiment scores of different algorithms can be analyzed.

For the data preprocessing, we used KNIME software to preprocess and clean the textual data. For the text analysis, we examine the perfume product dataset to find the most frequent words that have occurred in the reviews using Lancaster and Porter libraries in Python as well as N-gram analysis for the positive and negative reviews. For the sentiment analysis, we apply three different sentiment scoring algorithms/implementations, namely the VADER Python library [11], TextBlob Python library [12], and KNIME lexicon-based algorithm [13]. For the model evaluation, we use the random forest machine learning method to compare the prediction performance of the three methods. And lastly, we conduct a gap analysis between the VADER and TextBlob sentiment results and examine which factors are related to gaps in the sentiment scores of the selected two algorithms. These algorithms/implementations were selected due to their popularity in both academic literature and in business practice.

Two critical aspects/steps of the methodology are (a) the prediction of sentiment scores obtained by different algorithms and (b) analysis of gap between the sentiment scores obtained by different algorithms. To conduct (a), a new Dataset C was created, that combines Dataset B, which includes sentiment scores obtained through various algorithms, together with the data of term frequencies in each document.

4 Data

The original Amazon product reviews' dataset, which we refer to as Dataset A, was collected from Amazon.com by Ishani Chatterjee [4] and publicly shared online [3]. The data are separated into seven different CSV files, where each file includes data for a different product, Perfume, Book, Mask, Movie, Food, Curcumin, and Electronic. The reviews in each dataset were created between 2008 and 2020, and each of them has a collection of 5000 reviews and eight attributes. Each row in the dataset includes a review from an individual customer as well as additional review information such as ratings. Dataset A lists and describes the attributes included in the datasets, Product name, Ratings, Reviews, Helpful, Date, Asin, Target, and Text.

Using this source Dataset A, after excluding irrelevant attributes, appending sentiment scores, discretizing sentiment scores, and deriving new attributes (especially for cumulative values), a new Dataset B was obtained. While Dataset B has many attributes, the scope of the current study was limited to only some of these attributes, as a first step. One of the future research possibilities is to enrich and extend the current methodology to become much more comprehensive, yielding much richer insights by design. Still, the full meta-data for Dataset B is provided in Table 1 in this paper, to lay the foundation for future studies, as well as motivate other researchers to work with this readily enriched dataset.

5 Analysis and Results

5.1 Data Preprocessing

For data preprocessing, we apply a text analytics workflow within the KNIME platform to simplify the analysis and make the textual data ready for sentiment analysis without any noisy words or text errors. Some of the different KNIME nodes that were used in the text preprocessing are Case Convertor, Punctuation Erasure, Stope Word Filter, Dictionary Filter, N Chars Filter, and Number Filter. In addition, we have also filtered the infrequent terms that have occurred less than 10 times by using the Bag of Words (BoW) and GroupBy nodes.

Feature Selection and Engineering. Since we are only interested in the customer reviews and their associated ratings, we dropped many of the unrelated attributes (Product name, ASIN, Target, and Text) from the datasets and developed new variables that could create informative insights. The sentiment is determined by the customer's rating based on a scale of 1–5 (5 being the most favorable). As we are using classification methods to classify customer reviews, these scores will need to be converted into two categories, namely 1 and 0. Ratings above and including 4 will be labeled as positive reviews "1." Ratings with a score of 3 and below will be labeled as negative reviews "0."

Other features have been added to the dataset that could be contributed to the analysis of the customer review data, such as Cumulative Average Rating, Word Count, which is the number of words in each review, Cumulative Sum of Word Count, Character Count, Date Gap, and Day Since Last Review.

5.2 Text Mining

Most Frequent Words. Figure 2 displays the most frequent terms occurring in the customer's reviews for the perfume product. The word cloud has been generated by deleting stop words, such as "that," "the," and pronouns, as well as frequent words like "perfume," "product," and "amazon," which naturally occur in big portion of the reviews.

It is observed immediately from Fig. 2 that the term set retrieved from the reviews is mostly positive, which makes sense since the number of positive reviews is much higher than non-positive reviews. Most reviews/comments discuss the characteristics of the perfume, like the smell, scent, how long it lasts, etc. There are also terms indicating the feelings of customers, such as love and like.

Table 1 Shows the attributes in Dataset B, other types, and brief descriptions

No.	Attribute	Type	Description
1	RowID	Numerical	Unique ID for each review (each row is a review)
2	ProductNumber	Numerical	Number of each review for each product
3	ASIN	Numerical	Amazon Standard Identification Number
4	ProductName	Categorical	Name of the product
5	Ratings	Numerical	Rating of the product in that review: 1–5 (Likert scale)
6	RatingClass	Binary	Rating class; positive rating (1): 4–5; negative rating (0): 1–3
7	Review	Text	Customers' review
8	WordCount	Numerical	Number of words in each review
9	CharacterCount	Numerical	Number of characters in each review
10	Helpful	Numerical	How helpful the review is for other customers
11	Date	Date type	Date of the review
12	Year	Numerical	Year of when the review was written
13	Month	Numerical	Month of when the review was written
14	Day	Numerical	Day of when the review was written
15	DateCode	Numerical	Unique code of date of the review
16	DateGap	Numerical	Number of gap days from the first review to the date of this review
17	DaysSinceLastReview	Numerical	Number of days past since the last review
18	Target	Categorical	Targeted reviews: positive or negative
19	ProductType	Categorical	Product type/category (Food, Books, Masks, Perfume, Curcumin, Electronics, Movies)
20	CumulAvgRating	Numerical	Average of all the ratings until this review, excluding this review
21	CumulSumHelpful	Numerical	Summation of helpful values for all reviews until now
22	CumulSumWordCount	Numerical	Summation of Word Counts of all reviews until now
23	CleanedReview	Text	Review text after text cleaning and preprocessing
24	ScorePythonVADER	Numerical	Sentiment score of the Python library VADER [−1, 1]

(continued)

Table 1 (continued)

No.	Attribute	Type	Description
25	SentimentPythonVADER	Numerical	Sentiment label computed in VADER library of Python, for the text in Review {0, 1}
26	ScorePythonTextBlob	Numerical	Sentiment score of the Python library TextBlob [−1, 1]
27	SentimentPythonTextBlob	Numerical	Sentiment label computed in TextBlob library of Python, for the text in Review {0,1}
28	KNIMENegativeWords	Numerical	Number of negative words from the preprocessed review
29	KNIMEPositiveWords	Numerical	Number of positive words from the preprocessed review
30	WordCountCleanedReview	Numerical	Number of words in the preprocessed review
31	SentimentScoreKNIME	Numerical	Sentiment score of KNIME [−1,1]
32	SentimentKNIME	Numerical	Sentiment label computed in KNIME for the text in Review {0, 1}

Fig. 2 Word cloud of the perfume product most frequent words that have occurred in the customer reviews

5.3 *Sentiment Scoring*

Sentiment analysis is one of the most common core areas where NLP has been used. Businesses need to know how customers act and what they expect from the products and services they buy. Sentiment scoring is the process of assigning a score to a text document, usually between -1.0 and 1.0, where a score of 1.0 indicates a very positive sentiment and a score of -1.0 indicates a very negative sentiment. The outcome of the score is calculated based on the number of positive and negative words in the document and the way those words are used (e.g., whether they are used in a positive or negative context). The feedback a customer gives about a product can be "positive" or "negative." Interpreting feedback from customers through ratings and reviews enables businesses to measure customer satisfaction with their products and services. In addition to analyzing the polarity of a text, it can also identify certain sentiments and emotions, such as anger, happiness, and sadness. Even intentions, such as whether a person is interested or not, may be deduced using sentiment analysis.

For the sentiment analysis, several sentiment methods were conducted using the Python programming language and the KNIME platform. Our goal was to compare the performance of each method and find out which one of them has the most accurate performance in classifying customer reviews.

For the Python programming language, VADER and TextBlob sentiment analysis libraries were conducted using the Natural Language Processing (NLTK) package to determine the text's mood.

VADER Sentiment Analysis. Valence Aware Dictionary and Sentiment Reasoner (VADER), which is a rule-based and lexicon-based pre-built library in NLTK, is one of the best choices for sentiment analysis in Python. This library, which was developed particularly for social media sentiment analysis [11], includes a sentiment lexicon and a collection of lexical properties that are commonly categorized according to their sentiment polarity as either positive or negative.

VADER computes text sentiment and returns the likelihood that a particular input statement is positive, negative, or neutral. The library returns a compound score, which is also known as a polarity score, which is a measure that calculates the total of all normalized lexicon ratings between -1 (extremely negative) and $+1$ (extremely positive). To label the sentiment scores as positive and negative, we have classified the polarity scores as positive sentiment (polarity score > 0) and negative sentiment (polarity score ≤ 0).

TextBlob Sentiment Analysis. TextBlob is an ideal substitute for sentiment analysis. The basic Python library provides extensive textual data analysis and processing. TextBlob defines a sentence's mood based on its sentiment polarity and the intensity of each word, which requires a predefined dictionary to distinguish negative and positive terms. The tool gives each word a separate score and calculates what the overall emotion is [12]. TextBlob returns a sentence's polarity and subjectivity, with polarity ranging from negative to positive.

KNIME Sentiment Analysis. As part of KNIME's text processing feature, textual data were read, processed, and transformed into numerical data (documents and term vectors) to be used in regular KNIME data mining nodes for classification [13]. KNIME can analyze and parse texts in different formats and store the results in a table. In this way, the document can be further semantically enhanced by recognizing and tagging different kinds of named entities, such as those with positive and negative sentiments. Documents can be filtered in many ways, such as by using stop words nodes or named entities, stemming with stemmers that work with more than one language, and preprocessing in many different ways. Furthermore, it is possible to compute the frequency of words, extract keywords, and do some visualization in KNIME. Based on the document sentiment results, one can apply regular KNIME nodes to classify documents using numerical vectors. In this paper, we used the MPQA subjectivity lexicon [14] to identify contextual polarity depending on the lexicon-based approach.

5.4 N-Grams

N-Grams are combinations of “n” words within a sentence that can play an important part in text categorization and language modeling. In this analysis, the “ngram” method in the NLTK library is used to discover all n-grams in the review texts.

N-Grams for Positive Reviews. Using the N-Gram Python method ranging from 1 to 3 g, we have divided the positive and negative VADER sentiment score results separately to see what are the terms that are most repeated in each of the sentiment reviews. The results of the N-Grams for the VADER-positive sentiments for the perfume product indicated that most of the customers feel good about purchasing the perfume, where the number of times the words great and good have occurred more than 1000 times. Moreover, we observed that people also think about the scent as being fresh, lasting a long time, and smelling good and great as most of the written “smells great,” “love smell,” “fresh scent,” and “smells pretty good.” As well as the price of the product, where some of them have written that the price of the perfume is great such as “great price.”

N-Grams for Negative Reviews. These negative reviews of the perfume product, people who wrote these negative reviews believe the product is fake or smells bad and these words have occurred more than 50 times. Others have also given negative reviews because they have received broken perfume. But compared to the number of these terms that have occurred, it is not comparable to the number of the positive terms and how much they have occurred.

6 Sentiment Prediction

For the comparison of sentiment analysis between the three methods (VADER, TextBlob, and Lexicon-based algorithm), we labeled the sentiment scores for each product’s reviews as either positive or negative. Text mining was carried out to identify the most frequent terms, which in turn were considered as predictive features/attributes (columns in a tabular dataset) whose term frequency values were used for sentiment prediction. The classification algorithm used was the random forest machine learning algorithm, which enabled the comparison of the predictability of sentiments from the three sentiment scoring lexicons/algorithms. We aimed to assess

Table 2 Sentiment analysis methods' accuracy comparison

Products	VADER sentiment method	TextBlob sentiment method	KNIME sentiment method
Perfume	**0.938**	0.921	0.915
Books	0.886	0.888	**0.906**
Curcumin	0.898	**0.902**	0.865
Electronics	0.887	0.885	**0.906**
Food	0.893	0.891	**0.928**
Masks	**0.907**	0.905	0.904
Movies	0.874	**0.878**	0.865

Bold indicates the method that performs best

which machine learning algorithm model has the highest accuracy in predicting the sentiment of a customer review for each of these three algorithms.

6.1 Reviews' Sentiment Prediction Accuracy Comparison

Table 2 displays the sentiment prediction accuracy results using random forest classification. By looking at the accuracy of the product review sentiment prediction, we can figure out which algorithm is the most accurate and works best. As it is shown below, VADER algorithms achieved the highest accuracy for the perfumes and masks datasets. However, TextBlob algorithms performed well for the curcumin and movie datasets, while KNIME algorithms worked accurately with books, electronics, and food datasets. This shows that the three methods are good predictors for sentiment analysis since classification accuracies for all three methods are close to each other. However, to obtain the highest classification accuracy for different products, all three algorithms can be considered.

6.2 Sentiment Score Gap Analysis

The sentiment score gap analysis was conducted to find the difference between the results of the VADER and the TextBlob sentiment scores. By calculating the gap between the sentiment score results of the two methods (VADER score minus TextBlob score), we first analyze the correlation between the sentiment gap (y axis) and the other continuous variables for all the products (x axis). The results of the correlation suggested that in most of the products, the sentiment gap has a positive correlation with the word and character count of the review. Figure 3 depicts as scatter plots, the relationship between the sentiment gap (y axis) and the Word Count of the review (x axis). As we look into the scatter plots for the seven selected products, we

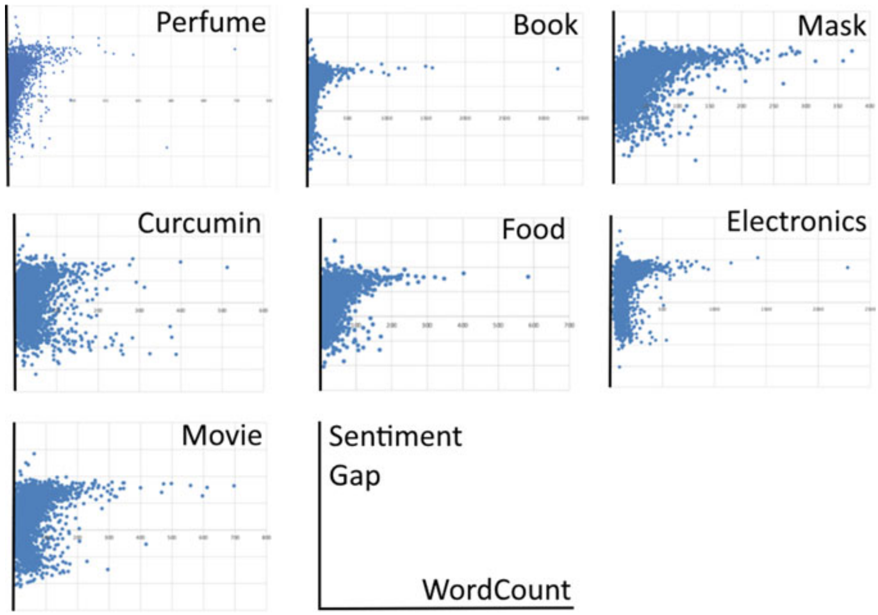


Fig. 3 Scatter plots of the relationship between the sentiment gap of VADER and TextBlob sentiment scores and the Word Count of the review for all the products

can notice a consistent patterns: As the number of words in the review increases, it is more likely to be labeled as a positive review by VADER, compared to TextBlob. Furthermore, the number of positive reviews for all the products is much higher than the number of negative reviews. There are many other analyses that have been and can be conducted, yet the content of this paper was kept limited to the analysis of only one gap analysis relation, due to the paper's space limitations.

7 Conclusion and Future Work

In this paper, we primarily focused on sentiment mining basics and their levels. The identification of sentiment from content can be achieved in several different ways. Sentiment analysis analyzes people's sentiments, attitudes, and emotions toward certain entities. In this paper, we addressed sentiment polarity categorization as a fundamental problem in sentiment analysis, which we focused on by categorizing customer/user opinions on select Amazon products as positive or negative. Furthermore, we studied the differences between three different sentiment algorithms (VADER, TextBlob, and KNIME).

References

1. Mehta P, Pandya S (2020) A review on sentiment analysis methodologies, practices and applications. Int J Sci Technol Res 9(2):601–609
2. Oxford languages. https://languages.oup.com/. Last accessed 19 Oct 2022
3. Harvard Dataverse. https://doi.org/10.7910/DVN/W96OFO. Last accessed 14 2022 Sept
4. Chatterjee I, Zhou M, Abusorrah A, Sedraoui K, Alabdulwahab A (2021) Statistics-based outlier detection and correction method for amazon customer reviews. Entropy 23(12):1645. https://doi.org/10.3390/e23121645
5. Fang X, Zhan J (2015) Sentiment analysis using product review data. J Big Data 2(1):1–4
6. Naseem U, Razzak I, Khushi M, Eklund PW, Kim J (2021) COVIDSenti: a large-scale benchmark Twitter data set for COVID-19 sentiment analysis. IEEE Trans Comput Soc Syst 8(4):1003–1015
7. Onan A (2021) Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. Concurrency Comput: Pract Exper 33(23):e5909
8. Rezaeinia SM, Rahmani R, Ghodsi A, Veisi H (2019) Sentiment analysis based on improved pre-trained word embeddings. Expert Syst Appl 1(117):139–147
9. Mowlaei ME, Abadeh MS, Keshavarz H (2020) Aspect-based sentiment analysis using adaptive aspect-based lexicons. Expert Syst Appl 15(148):113234
10. Al-Shabi MA (2020) Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions Mining. IJCSNS. 20(1):1
11. vaderSentiment, https://pypi.org/project/vaderSentiment/. Last accessed 19 Oct 2022
12. TextBlob, https://textblob.readthedocs.io/en/dev/. Last accessed 19 Oct 2022
13. Bessa A (2022) Lexicon-based sentiment analysis: A Tutorial, 2022/03/17 https://www.knime.com/blog/lexicon-based-sentiment-analysis. Last accessed 141 Sept 2022
14. MPQA Opinion Corpus Release Page, https://mpqa.cs.pitt.edu/corpora/mpqa_corpus/. Last accessed 141 Sept 2022

Formal Stability Analysis of Two-Dimensional Digital Image Processing Filters

Adnan Rashid, Sa'ed Abed, and Osman Hasan

Abstract There are several image processing applications that require to partition frequency components of images. This requirement is usually fulfilled by using digital image processing filters. Most of this processing is done in two-dimensions given the two-dimensions of the regular images. It is very important to ascertain that these filters provide a stable output for a bounded input, and this requirement is usually termed as stability. The stability analysis of these filters is usually conducted analytically, on a piece of paper, or by simulations. However, these techniques provide approximate or inaccurate results as paper-based analysis can have human error and simulations suffer from computer arithmetic related roundoff limitations. We advocate formally analyzing the stability of digital filters for two-dimensional (2D) images using interactive theorem proving. In this regard, we present a formal dynamical model and a formal notion of stability of 2D digital image processing filters in HOL Light. The proposed formal model is used to perform the stability analysis of a real-world 2nd-order filter in HOL Light.

Keywords Stability · 2D z-transform · Interactive theorem prover

A. Rashid (✉) · O. Hasan
School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan
e-mail: adnan.rashid@seecs.nust.edu.pk

O. Hasan
e-mail: osman.hasan@seecs.nust.edu.pk

S. Abed
Department of Computer Engineering, College of Engineering and Petroleum, Kuwait University, Kuwait City, Kuwait
e-mail: s.abed@ku.edu.kw

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_47

1 Introduction

Digital image processing filters (IPFs) are extensively being used in many application areas, such as medicine [1] and autonomous vehicles [2, 9], for performing different operations, like image processing, filtering and enhancement, in two-dimensional (2D) images. For example, they are used to pre- and post-process images to filter elements like noise and image smoothing and quality enhancement by filtering out the noise and distortion [1]. Filters are mainly of three kinds, namely lows, band, and high pass. For example, a high-pass filter can be used for the passage of frequencies higher than a certain range.

Stability of a digital IPF asserts a stable output response to a given bounded input and is considered as an important phenomenon for accessing the performance of an IPF. For a 2D digital IPF, it is described in terms of the transfer function, i.e., the relationship of output to input in the frequency domain. To analyze the stability analysis of a 2D digital IPFs, we first need to capture their dynamical behavior in terms of 2D difference-equations (DEs). Next, the 2D z-transform is utilized for their analytical analysis by converting the DEs to algebraic equations, i.e., transforming the 2D arrays to the (z_1, z_2)-domain. Lastly, these (z_1, z_2)-domain representations are utilized for the stability analysis [13].

Conventionally, the stability analysis of the digital IPFs has been conducted using analytically on paper or simulations. But these methods, due to their human error proneness and round off errors, cannot guarantee accurate results. Therefore, these conventional approaches cannot be relied upon considering the wider utility of these filters in many critical domains, like transportation and healthcare.

Formal verification [8] is an analysis approach that involves capturing the behavior of the given system in the form of a logical model and verifying the system characteristics deductively in a computer. Interactive theorem proving [4, 7] is one of the extensively used formal verification techniques. We argue to use interactive theorem proving to conduct the stability analysis of the digital IPFs. In this regard, we formalize the dynamics of the digital IPF as a 2D array in the HOL Light prover [6] . This model is then used to perform the formal stability analysis based on the transfer function, obtained using the z-transform on the dynamical model of the digital filter. We chose HOL Light for our work as it has a strong reasoning support for multivariate calculus and digital IPFs [12]. These existing works have greatly facilitated our formalization as we built upon them to develop reasoning support for the stability analysis of digital IPF.

We introduce HOL Light and some of definitions and some of the utilized theorems from HOL Light's theory of the multivariable calculus in Sect. 2. Section 3 describes the modeling of 2D z-transform in HOL Light. The formal model for stability of the 2D digital IPFs in presented in Sect. 4. The formal stability analysis of the 2nd-order digital IPF is described in Sect. 5. Finally, Sect. 6 provides some insights that we gathered from our work as well our plans to further extend out reasoning support.

2 Preliminaries

We present some background information in this section to help the reader in understanding the remaining paper.

2.1 Interactive Theorem Prover: HOL Light

HOL Light [5], developed using ML [11], is an interactive proof-assistant that is widely used for developing proofs for the mathematical concepts and analyzing software and hardware systems. A theorem is a mathematical statement that can be proved using a predefined set of primitive rules or axioms in a theorem prover, ensuring the soundness of the proof development environment. HOL Light contains several multivariate theories, in particular, vectors, differential, integral, and 2D z-transform, which are used in the proposed work.

2.2 Multivariable Calculus

A generic vector is modeled as a N element matrix, i.e., $\mathbb{R}^N$, of real numbers. This model allows us to use matrix operations for vector manipulations.

Summation over a generalized function f of an arbitrary datatype A $\rightarrow \mathbb{R}^N$ is modeled as

Definition 1 $\vdash_{def}$ ∀s f. vc_smm s f = (lambda j. smm s (λy. f y$j))

where vc_smm accepts an arbitrary set s: and a function f as inputs and outputs the vector-addition on s, which represents a set. smm models a finite summation over f and thus vc_smm (0..n) f mathematically models $\sum_{j=0}^{n} f(j)$. Similarly, we formalize the mathematical expression $\sum_{j=0}^{\infty} f(j) = l$, involving an infinite summation for a function f of datatype $\mathbb{N} \rightarrow \mathbb{R}^N$ and a limit value l of datatype $\mathbb{R}^N$, in HOL Light as follows:

Definition 2 $\vdash_{def}$ ∀s f l. (f smms l) s ⇔
((λn. vc_smm (s ∩ (0..n)) f) → l) squntialy

where squntialy mathematically represents a sequential growth, i.e., $f(j)$, $f(j+1)$, ..., etc.

Definition 3 $\vdash_{def}$ ∀f s. smmble f s ⇔ (∃l. (f smms l) s)

The HOL Light function smmble mathematically models $\sum_{j=0}^{\infty} f(j) = l$.

Next, we present the formal modeling of the infinite summation:

Definition 4 $\vdash_{def} \forall \mathsf{f}\ \mathsf{s}.\ \mathsf{inft\_smm\ s\ f} = (\in \mathsf{l}.\ (\mathsf{f\ smms\ l})\ \mathsf{s})$

where the return value l: $\mathbb{R}^N$ is the value of the infinite summation of the converging function f from the given starting point s.

3 Formal Modeling of the 2D z-Transform

The z-transform of a 2D discrete-time function $f(m_1, m_2)$ is expressed as [13]:

$$F(z_1, z_2) = \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} f(m_1, m_2) {z_1}^{-m_1} {z_2}^{-m_2} \tag{1}$$

We formalize Eq. (1) as follows:

Definition 5 $\vdash_{def} \forall \mathsf{f}\ z_1\ z_2.\ \mathsf{z\_2d\_trnsfm\ f}\ z_1\ z_2 = \mathsf{inft\_smm\ (from\ 0)}$
$$\left(\lambda m_1.\ \mathsf{inft\_smm\ \ (from\ \ 0)}\ \left(\lambda m_2.\ \frac{\mathsf{f}\ m_1\ m_2}{{z_1}^{m_1} * {z_2}^{m_2}}\right)\right)$$

Here, we need to identify the set of all values of z_1 and z_2 for which the infinite summations converge to some finite value and thus ensure a finite $F(z_1, z_2)$, commonly known as the region of convergence (ROC). We can mathematically express and formally model the ROC as follows:

$$\mathrm{ROC} = z_1, z_2 \in \mathbb{C} : \exists k.\ \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} f(m_1, m_2) {z_1}^{-m_1} {z_2}^{-m_2} = k \tag{2}$$

Definition 6 $\vdash_{def} \forall \mathsf{f}\ m_1.\ \mathsf{z\_2d\_ROC\ f}\ m_1 =$
$\{(z_1, z_2) \mid (z_1 \neq 0) \wedge (z_2 \neq 0) \wedge$
$\mathsf{z\_2d\_tr\_smmble\ f}\ z_1\ z_2\ m_1 \wedge \mathsf{z\_2d\_tr\_td\_smmble\ f}\ z_1\ z_2\}$

where z_2d_ROC takes a function f and m_1, which represents the starting point in Eq. (1), as inputs and outputs a set of non-zero values of variables z_1 and z_2 for which the 2D z-transform of f exists. We also formalized functions z_2d_tr_smmble and z_2d_tr_td_smmble capturing the summability of the function f for the inner and the outer (double) summations, respectively as

Definition 7 $\vdash_{def} \forall \mathsf{f}\ z_1\ z_2\ m_1.\ \mathsf{z\_2d\_tr\_smmble\ f}\ z_1\ z_2\ m_1 =$
$$\left(\forall m_1.\ \mathsf{smmble\ \ (from\ \ 0)}\ \left(\lambda m_2.\ \frac{\mathsf{f}\ m_1\ m_2}{{z_1}^{m_1} * {z_2}^{m_2}}\right)\right)$$

Definition 8 $\vdash_{def} \forall \mathsf{f}\ \mathsf{z1}\ \mathsf{z2}.\ \mathsf{z\_tr\_td\_summable\ f\ z1\ z2} = \mathsf{summable\ (from\ 0)}$
$$\left(\lambda m_1.\ \mathsf{inft\_smm\ \ (from\ \ 0)}\ \left(\lambda m_2.\ \frac{\mathsf{f}\ m_1\ m_2}{{z_1}^{m_1} * {z_2}^{m_2}}\right)\right)$$

Now, we formally verify some key characteristics of the 2D z-transform, including linearity, shifting, scaling, complex conjugation, and 2D z-transform of a n-order system, in HOL Light. This formally verified characteristics play a key role in the proposed stability analysis of the 2D digital IPFs, as presented in Sect. 5. The 2D z-transform, presented in this section, has been formalized by Rashid et al. [12]. However, the authors have not performed the stability analysis of the 2D digital IPF, which is indeed the scope of this paper. The actual formalization of the 2D z-transform can be viewed at.[1]

4 Stability of a 2D Digital Image Processing System

Stability is considered as an important characteristic while designing a 2D digital IPF. A discrete-time system such as a digital filter is said to be stable if it provides a bounded output for a given bounded input. An important condition for the stability of a linear shift invariant (LSI) system can be mathematically expressed as [10]:

$$\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} |h(m_1, m_2)| < \infty \tag{3}$$

where $h(n_1, n_2)$ provides the impulse response, i.e., the output response when the input is a brief input function of the given LSI system. However, it is more convenient to represent stability based on the system function/transfer function $H(z_1, z_2)$ (the Laplace transform of $h(m_1, m_2)$), which is mathematically expressed as

$$H(z_1, z_2) = \frac{Y(z_1, z_2)}{X(z_1, z_2)} \tag{4}$$

According to Shanks, the stability of a LSI system such as digital filter can be mathematically expressed by the two conditions as follows [10]:

$$\begin{aligned} &\text{Stability} \Leftrightarrow \quad (a)\ X(z_1, z_2) \neq 0 \ \text{ for } \ |z_1| = 1, |z_2| \geq 1 \\ &\text{and} \ \ (b)\ X(z_1, z_2) \neq 0 \ \text{ for } \ |z_1| \geq 1, |z_2| = 1 \end{aligned} \tag{5}$$

We can use the following two steps to ensure Condition (a) of the stability for the digital IPF. In the first step, we need to solve for all (z_1, z_2), such that $X(|z_1| = 1, z_2) = 0$, which is equivalent to solving for all (ω_1, z_2), such that $X(e^{j\omega_1}, z_2) = 0$. In the next step, we have to check that if all $|z_2|$ obtained in the first step are less than 1. Similarly, we can use the similar steps to ensure Condition 2 of the stability. Using this alternative representation, we formalized the stability of a digital filter in HOL Light as follows:

[1] http://save.seecs.nust.edu.pk/fsadipf/.

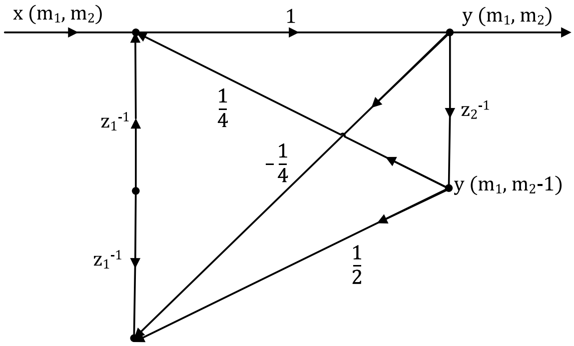


Fig. 1 Flowgraph of a 2nd-Order 2D IPF

Definition 9 $\vdash_{def}$ ∀X. cnd1_stbl_dgtl_fltr X =
{(ω_1, z_2) | X ($e^{j\omega_1}$, z_2) = 0 ∧ |z_2| < 1} ≠ { }
$\vdash_{def}$ ∀X. cnd2_stbl_dgtl_fltr X = {(z_1, ω_2) | X (z_1, $e^{j\omega_2}$) = 0 ∧ |z_1| < 1} ≠ { }
$\vdash_{def}$ ∀X. is_stbl_dgtl_fltr X = cnd1_stbl_dgtl_fltr X ∧ cnd2_stbl_dgtl_fltr X

where is_stbl_dgtl_fltr accepts the denominator X of the transfer function, provided in Eq. (4), corresponding to the dynamics of a digital IPF and provides a stable filter.

5 Formal Stability Analysis of a 2nd-order Filter

We utilize the formalization, provided in Sects. 3 and 4, for performing the formal stability analysis of a 2nd-order 2D digital IPF in this section. This illustrates the practical utilization of the foundational formal modeling, presented in this paper.

Graphically, we can present a 2nd-order 2D digital IPF by the flowgraph depicted in Fig. 1. It is a collection of nodes and branches, which provide the directed connections between these nodes. The constants 1, $\frac{1}{4}$, $-\frac{1}{4}$, and $\frac{1}{2}$ in Fig. 1 present the gains of each branches. Similarly, ${z_1}^{-1}$ and ${z_2}^{-1}$ model the horizontal and vertical delay, i.e., shift right and shift up, operations, respectively. This 2nd-order digital IPF can be mathematically expressed using the following linear difference equation (DE).

$$\begin{aligned} y(m_1, m_2) = x(m_1, m_2) + \frac{1}{4}y(m_1, m_2 - 1) - \frac{1}{4}y(m_1 - 2, m_2) \\ + \frac{1}{2}y(m_1 - 2, m_2 - 1) \end{aligned} \tag{6}$$

We can mathematically describe the transfer function of the 2nd-order digital IPF corresponding to its dynamical model [Eq. (6)] as follows:

$$H(z_1, z_2) = \frac{Y(z_1, z_2)}{X(z_1, z_2)} = \frac{1}{1 - \frac{1}{4}{z_2}^{-1} + \frac{1}{4}{z_1}^{-2} - \frac{1}{2}{z_1}^{-2}{z_2}^{-1}} \tag{7}$$

The main purpose of presentation this case study is to use our proposed formal models to formally verify Eq. (7). To verify the transfer function, the first step is to formally model the DE of the filter [Eq. (6)] as follows:

Definition 10 $\vdash_{def}$ ∀y x m1 m2 p q. dgtl_scnd_odr_fltr x y p q m1 m2 ⇔
y (m1, m2) = l1l2th_dfrnce_eq y p 2 2 m1 m2 - l1l2th_dfrnce_eq x q 0 0 m1 m2

with coefficients a and b of the input and output 2D arrays. The function dgtl_scnd_odr_fltr accepts the 2D arrays x and y, their coefficients a and b, and it uses the (L_1, L_2)-order DE l1l2th_dfrnce_eq to capture the linear DE expressing the 2nd-order digital IPF.

We verify Eq. (7) as follows:

Theorem 1 $\vdash_{thm}$ ∀x y p q z_1 z_2 m_1.
[C_1]: (z_1, z_2) IN 2d_roc_lccdifeq x y 2 2 q m_1 $\wedge$
[C_2]: in_frst_qudrnt_2d_lccdifeq x y $\wedge$
[C_3]: $z_1 \neq 0 \wedge$ [C_4]: $z_2 \neq 0 \wedge$
[C_5]: (∀m_1 m_2. dgtl_scnd_odr_fltr x y p q m_1 m_2)

$$\Rightarrow \frac{\text{z\_2d\_trnsfm y } z_1\ z_2}{\text{z\_2d\_trnsfm x } z_1\ z_2} = \frac{1}{1 - \frac{1}{4} * z_2{}^{-1} + \frac{1}{4} * z_1{}^{-2} - \frac{1}{2} * z_1{}^{-2} * z_2{}^{-1}}$$

Condition C_1 provides the ROC of the dynamical model of the 2nd-order digital IPF. Condition C_2 ensures the first quadrant conditions on the input (x) and output (y) 2D arrays. Conditions C_3 and C_4 assert the non-zero condition for the variables z_1 and z_2. Condition C_5 presents the dynamical model of the 2nd-order digital filter captured by Eq. (6). The transfer function of the IPF is verified based on these assumptions as the conclusion of the theorem. The verification of Theorem 1 depends on the formal development of the 2D z-transform described in Sect. 3.

Next, we use the transfer function to formally verify the stability of the 2nd-order 2D digital IPF as follows:

Theorem 2 $\vdash_{thm}$ ∀z_1 z_2. [C_1]: $z_1 \neq 0 \wedge$ [C_2]: $z_2 \neq 0 \wedge$

$$\Rightarrow \text{is\_stbl\_dgtl\_fltr} \left(\frac{1}{1 - \frac{1}{4} * z_2{}^{-1} + \frac{1}{4} * z_1{}^{-2} - \frac{1}{2} * z_1{}^{-2} * z_2{}^{-1}} \right)$$

Conditions C_1 and C_2 assert the non-zero condition for the variables z_1 and z_2. Finally, the conclusion models the stable 2nd-order IPF. The verification of the above theorem is based on formalization of the stability, provided in Sect. 4.

Finally, we implement Condition (a) of the stability of the 2nd-order digital IPF (Theorem 2) using the Python language. For this, we implement the characteristic equation $1 - \frac{1}{4}z_2{}^{-1} + \frac{1}{4}z_1{}^{-2} - \frac{1}{2}z_1{}^{-2}z_2{}^{-1} = 0$ on the complex plane z_2 for $z_1 = e^{i\omega_1}$, $\omega_1 \in [0, \pi]$. In the case of the 2nd-order digital IPF (Fig. 2), the presence of poles inside the unit circle contributes to the stability of the filter. Similarly, we can

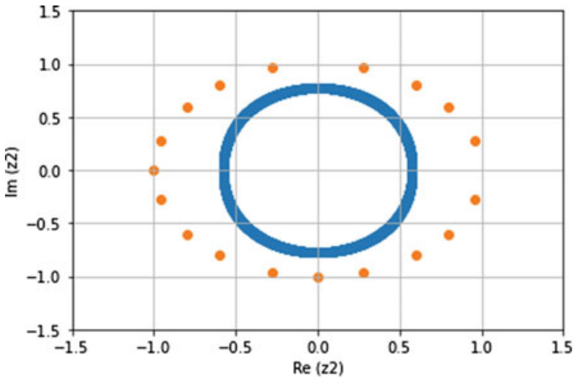


Fig. 2 Stability of the 2nd-order digital IPF on root map

implement Condition (b), which alongside Condition (a) ensure the stability of the corresponding filter.

The main novelty of our results is that the generic nature of the verified properties, i.e., all theorems are verified for the universally quantified variables and functions. For example, we have formalized the dynamical model of the 2nd-order filter using the (L_1, L_2)-order linear differential equations by specializing the generalized gains $(\alpha(l_1, l_2), \beta(k_1, k_2)$ to some particular values. Another positive aspect of our formal stability analysis, presented in this paper, is the assurance of explicit presence of all the required assumptions along with the theorem that are often ignored in the traditional methods. These advantages are obtained at the cost of significant involvement of a user in the formal stability analysis, due to the usage of an interactive theorem proving tool. To reduce this user intervention, we proposed several simplifiers such as DFRNC_EQU_TAC and TRNSFR_FNCTN_TAC[2] that significantly reduce the user guidance in the reasoning process.

6 Conclusions

Stability of a digital IPF is one of their important characteristics ensuring a stable output for a bounded input. We advocate using interactive theorem proving for performing stability analysis of these filters. In this regard, we formalized a dynamical model of the digital IPF and used the 2D z-transform to formally conduct the stability analysis. Finally, as a case study, we performed the stability analysis of a 2D digital IPF. In future, we aim to model the 2D convolution to develop formal reasoning support for systems-of-systems involving various image processing tasks [3].

Acknowledgements This work was supported and funded by Kuwait University, Research Project No. (EO 07/19).

[2] https://save.seecs.nust.edu.pk/fsadipf/

References

1. Behrenbruch C, Petroudi S, Bond S, Declerck J, Leong F, Brady J (2004) Image filtering techniques for medical image post-processing: an overview. British J Radiol 77(2):S126–S132
2. Blasinski H, Farrell J, Lian T, Liu Z, Wandell B (2018) Optimizing image acquisition systems for autonomous driving. Electron Imaging 2018(5):1–161
3. Dudgeon DE (1983) Multidimensional digital signal processing. Engewood Cliffs
4. Gordon MJ (1988) HOL: a proof generating system for higher-order logic. In: VLSI specification, verification and synthesis, SECS, vol 35. Springer, pp 73–128
5. Harrison J (1996) HOL light: a tutorial introduction. In: Srivas M, Camilleri A (eds) Proceedings of the first international conference on formal methods in computer-aided design (FMCAD'96). Lecture Notes in Computer Science, vol 1166. Springer, pp 265–269
6. Harrison J (1996) HOL light: a tutorial introduction. In: Formal methods in computer-aided design, vol 1166. LNCS, Springer, pp 265–269
7. Harrison J (2009) Handbook of practical logic and automated reasoning. Cambridge University Press
8. Hasan O, Tahar S (2015) Formal Verif Methods. Encyclopedia of information science and technology, IGI Global Pub, pp 7162–7170
9. Hussain R, Zeadally S (2018) Autonomous cars: research results, issues, and future challenges. IEEE Commun Surv Tutor 21(2):1275–1313
10. Lim JS (1990) Two-dimensional signal and image processing. Englewood Cliffs
11. Paulson L (1996) ML for the working programmer. Cambridge University Press
12. Rashid A, Abed S, Hasan O (2022) Formal analysis of 2D image processing filters using higher-order logic theorem proving. EURASIP J Adv Sign Process 2022(1):1–18
13. Woods JW (2006) Multidimensional signal, image, and video processing and coding. Elsevier

Development of a Web-Based Strategic Management Expert System Using Knowledge Graphs

İlter İrdesel, Gurdal Ertek, Ahmet Demirelli, Lakshmi Kailas, Ahmet Lekesiz, and Riaz Uddin Shuvo

Abstract In this paper, we present the development of a Web-based expert system, *StrategyAdvisor Cloud*, to support strategic management decision-making. The system was developed using a multistage methodology that builds upon knowledge graphs, where knowledge acquisition and rule base construction by project members with different roles, capabilities, and skills can be facilitated through customized visual languages. The methodology systematizes knowledge acquisition and knowledge representation for each stage, coupled with algorithms for the transformation of knowledge graphs between successive stages. The developed expert system and the development process are described in detail in the paper and its supplement, to serve as guidance in the development of similar systems in future.

This research was funded by the United Arab Emirates University, Startup Grant Funding 31B127.

İ. İrdesel
Magneti Marelli, Bursa, Türkiye
e-mail: ilter.irdesel@magnetimarelli.com

G. Ertek (✉)
College of Business and Economics, United Arab Emirates University, Al Ain, UAE
e-mail: gurdal@uaeu.ac.ae

A. Demirelli
Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul, Türkiye
e-mail: ahmetdemirelli@sabanciuniv.edu

L. Kailas
College of Business and Economics, United Arab Emirates University, Al Ain, UAE
e-mail: lakshmi.kailas@gmail.com

A. Lekesiz
Faculty of Engineering, Marmara University, Istanbul, Türkiye
e-mail: ahmetlekesiz@gmail.com

R. U. Shuvo
Code Optimizer, Dhaka, Bangladesh
e-mail: shuvo@codeoptimizer.net

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_48

Keywords Knowledge graphs · Knowledge representation · Decision support systems · Expert systems · Strategic management

1 Introduction

This paper reports the development of StrategyAdvisor Cloud[1] (Figs. 1, 2, and 3), an expert system in the domain of strategic management. The developed Web-based expert system acquires information through a series of diagnostic questions and makes strategic policy recommendations based on the answers.

An *expert system (ES)* can be defined as "a computer system that simulates the decision-making ability of a human expert" [5]. Expert systems have been used extensively to support decision-making in diverse domains, such as medical, military, chemistry, engineering, manufacturing, and management [17]. In expert system development, tacit knowledge from experts is extracted and encoded as explicit codified knowledge in a knowledge base.

A *rule-based expert system (RBES)* encapsulates expert knowledge as IF-THEN *rules*, also called *production rules*.

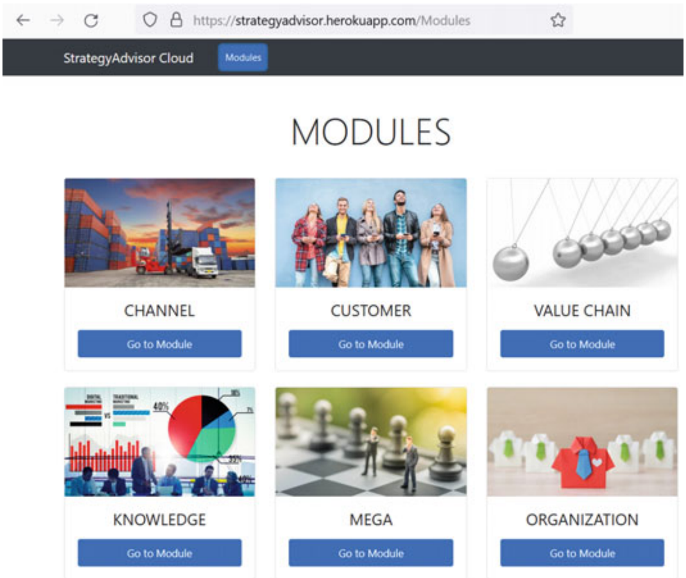


Fig. 1 Home modules page of StrategyAdvisor Cloud, from where the module of interest can be selected

[1] https://strategyadvisor.herokuapp.com/.

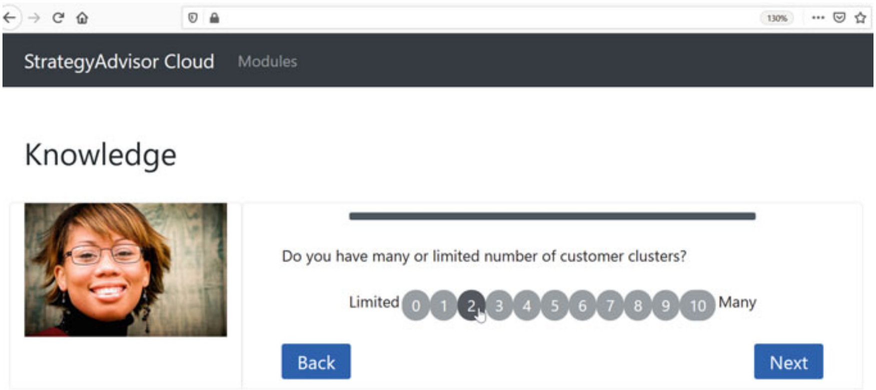


Fig. 2 Question page in StrategyAdvisor Cloud, which acquires facts through questions

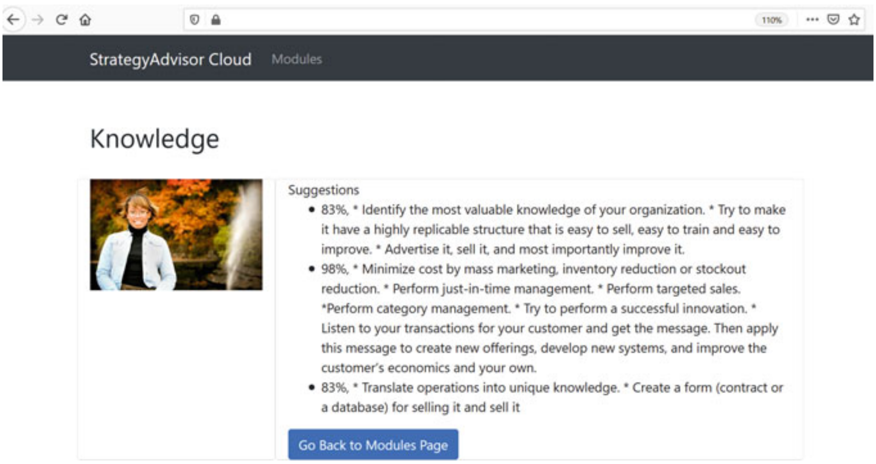


Fig. 3 Suggestion page in StrategyAdvisor Cloud, which suggests business strategies

RBES can help significantly in turning data and information into reusable and scalable knowledge assets, forming the engine of decision support systems (DSSs). The integration of rule engines into enterprise resource planning (ERP) systems, such as SAP BRM [9] and Oracle BPM [12], signals the increasing future adoption of rule-based systems for industry, business, and government applications.

This paper describes an expert system developed through the representation of knowledge in RBES through knowledge graphs. The applied multistage methodology supports the differentiated stages, processes, and team members in the development of expert systems. The motivation and objective are to facilitate knowledge acquisition and rule base construction by project members, each of whom has differing roles, tasks, priorities, capabilities, cognitive preferences, and technical competencies. Another primary motivation of the project was that there was no such Web-based system reported in the literature until now.

2 Background

This section provides a background on the challenges of developing expert systems and knowledge graphs as viable solutions. First, as the research motivation, challenges of knowledge acquisition in expert system development are discussed. Second, visualization and visual languages are identified as solution to the mentioned challenges. Third, the research gap in the literature, which the present research aims to fill, is identified and described.

2.1 Challenges of Knowledge Acquisition

Wagner et al. [18] reported that knowledge acquisition is the greatest bottleneck in the expert system development process because of the unavailability of experts and knowledge engineers, as well as the difficulties of the rule extraction process.

The fact that rules are eventually represented within the expert system detaches the domain expert from the knowledge representation, once the domain expert's tacit knowledge becomes codified explicitly as a text-based rule base. Furthermore, the expert system development process is not transparent to the domain expert and the eventual user, who typically have less technical rigor than business analysts and system designers. However, explicit rule representation *is* inevitably needed to process the knowledge by the rule engine, creating a dilemma. Similar problems may arise in the subsequent stages of the development process, where the agents may suffer from the cognitive overload of not being able to model or work with the constructs that match their roles, tasks, priorities, capabilities, cognitive preferences, and technical competencies. The mainstream expert system development environments support only one or two stages and roles, resulting in a mismatch between functional requirements and cognitive capabilities.

2.2 Visualization and Visual Languages

Visualization can be a feasible and viable solution for the development and implementation of management information systems (MIS), which also include DSS and ES. Kernbach et al. [8] suggest that graphically visualizing strategies helps managers to consider strategies better and even remember them further. The authors conducted an experiment with 76 managers to identify the impact of three types of visual formats on the effectiveness of strategy communication. Zabukovec and Jaklič [19] discussed the significance of modifying visualizations for different categories of users and situations to facilitate better handling of business data. Nissen [10] analyzed organizational knowledge through a system for visualizing and measuring it using knowledge flow principles.

While visualization has several advantages for knowledge assimilation, visual languages possess several advantages over text-based languages for software development [4]. Given such advantages, visual languages are adopted in the present study for expert system development, which visualize knowledge in different structures at different stages.

2.3 Research Gap

An extended review of the strategic management expert systems (SMES) literature revealed the incongruence between the technical capabilities of experts and existing knowledge acquisition methodologies. The review also revealed the incongruence between the expert systems development languages and the skills of the human agents (project team members and end-users) using them. While visualization can potentially help in knowledge acquisition, representation, and processing, a gap was identified in the research literature regarding the use of visualization and visual languages to resolve the mentioned incongruence in the selected domain of strategic management. Finally, a gap of well-documented know-how was identified regarding the development of Web-based expert systems for strategic management consulting.

3 Research Topic and Contributions to Literature

The topic of this research is the development of an expert system for strategic consulting using a multistage methodology (Fig. 4) built on knowledge graphs.

The contributions of this research are as follows: (1) The developed expert system, *StrategyAdvisor Cloud*, is a digital software as a service (SaaS) platform that performs strategic management consulting (Figs. 2 and 3). (2) The integrated multi-stage multi-agent methodology that is applied is used for the first time for strategic management domain. (3) For the first time in strategic management literature, visual representations suitable for each stage of the system development lifecycle are identified and formally specified through visual languages and mathematical notation. These representations are the *mind map*, *domain objects map* (DOM), and *rule map* (Figs. 6, 7, and 8, respectively). The visual representations are in accordance with the goals and tasks of that stage and the attributes of the agents in that stage. (4) For the first time in SMES literature, and possibly the larger expert systems literature, the transitions between the knowledge graphs of the successive stages are formally described as formal graph transformation algorithms. One of these algorithms, the transformation algorithm that transforms the Stage 1 mindmap into Stage 2 DOM is presented in 5 as an illustration. Other transformation algorithms are fully provided in the supplement [7] .

The idea of customized visual languages for different stakeholders in SMES was first introduced by İrdesel [6]. The cited work also includes the application of the

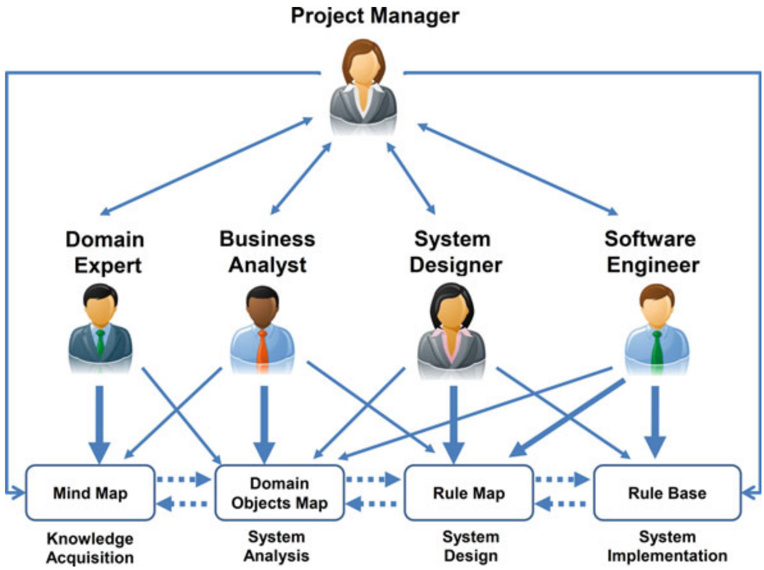


Fig. 4 Agents and the stages in expert systems development and the knowledge representations suggested for each stage by the methodology

idea for strategic management through an early version of the StrategyAdvisor desktop software, as well as the testing of the software through field studies involving more than 200 companies. However, the aforementioned work lacked a theoretical foundation, impeding its generalization and applicability in other applications within strategic management and the field of management at large, as well as in other diverse domains. The supplement [7] to our paper presents a strong theoretical foundation, where the knowledge graphs and the transformation between them are described through mathematical formalism. Graph transformations can be used for knowledge representation and verification, especially when the knowledge is dynamic [3]. Through the theoretical abstraction and foundation introduced in the current paper, it is possible to formally and methodologically apply the framework not only in strategic management but also in other domains. Transforming knowledge in any domain or application area into assets can be facilitated through the visual modeling of rule-based expert systems.

Contribution 1, which is the development of *StrategyAdvisor Cloud*, an expert system for selecting business strategy, is a case study application and illustration of the methodology. The system is developed through the acquisition of strategic management (SM) knowledge in the *Profit Patterns* book by Slywotzky et al. [14]. The selected book is structured such that it facilitates knowledge extraction and representation of tacit knowledge as explicit. While there are many more recent high quality books and other sources for strategic management, the strategies in the profit patterns book [14] are still strongly applicable after two decades, cementing

output G_2 **Transform_MindMap_to_DOM**(input G_1)
initial DOM is a replica of the mind map
$G_2(N_2, A_2) \leftarrow G_1(N_1, A_1)$
create a list of all paths that fire a suggestion
$P^* = \langle m, s, l, o, a_1, a_2, \ldots, v^* \rangle$
remove all arcs that emanate from logic nodes and terminate at object nodes
remove $\forall(l, o) \in A_2$
reverse the direction of all arcs from module nodes to logic nodes
$\forall(m, s) \in A_2$
    $(m, s) \leftarrow (s, m)$
reverse the direction of all arcs from suggestion nodes to logic nodes
$\forall(s, l) \in A_2$
    $(s, l) \leftarrow (l, s)$
an arc is drawn from every firing value to the suggestion that it fires
$\forall p^* = \langle m, s, l, o, a_1, a_2, \ldots, v^* \rangle \in P^*$
    $A_2 \sqcup (v^*, l)$
the transformation is complete, and the DOM can be returned as the output
return G_2

Fig. 5 Transformation algorithm that transforms the mindmap into DOM

the book as an evergreen classic of strategic management. Contribution 1 is novel in several ways: even though there have been other applications in the strategic management domain, the presented work is the only one that replicates the knowledge acquisition process in strategic management consulting. Similar to Surma [16], the present research builds on strategic patterns or cases. Thus, the case study described in this paper follows a "case-based patterns" approach [1, 13]. However, in contrast to earlier studies, the case study here is based on the cases and patterns formulated by a leading thinker in strategic management. To the best of our knowledge, this is the first study in which profit patterns formulated in the *Profit Patterns* book are codified as an expert system.

The novelties with contributions 2, 3, and 4 are explained in detail in the supplement [7].

The present study is the first in the strategic management literature with all four listed contributions. An extensive review of relevant research is provided in the supplement [7], where the study in this paper is compared to earlier related work.

4 Methodology

In this section, the stages of the applied graph-based methodology (Fig. 4) are described and discussed in further detail.

4.1 Overview

Figure 4 displays the steps of the methodology with reference to the human agents involved. The involvement of each agent at each stage is shown through arcs. The thicknesses of the arcs denote the level of involvement, with thicker arcs denoting a higher level of involvement. The round boxes are the knowledge representation schemes, and the texts below the boxes are the primary tasks at that stage. In Fig. 4, dashed arrows between the stages denote the translation of the rule base to the neighboring stages through graph transformation algorithms.

4.2 Agents and Stages

The applied methodology is agent-oriented; it caters to the tasks and goals of the human agents (project team members) involved in expert system development. These agents are *project managers, domain experts, business analysts, system designers,* and *software engineers*. The knowledge representation scheme for each stage (visual languages in the first three stages) is determined based on the focus of the primary team member active at that stage: In Stage 1, the task is knowledge acquisition, and mind map is suggested for this stage. The mind map first branches into the profit strategies, reflecting the focus of the domain expert (Fig. 6). In Stage 2, the task is system analysis and design, and domain objects map (DOM) is suggested. The DOM branches first into the domain objects, reflecting the focus of business analysts, who focus on identifying the elements in the system (Fig. 7). In Stage 3, the task is to model the logic for expert decision-making, and a rule map is suggested (Fig. 8). The rule map appeals to the system designer, who distinguishes between logic (rule base) and flow (rule engine) in designing the expert system. Finally, in Stage 4, the task is to transform the expert system into a stand-alone software or service, the principal task of the software engineer. Figure 9 illustrates a database structure that can support this stage, and Figs. 2 and 3 illustrate an example implementation.

4.3 Stages

The stages of the methodology are described in detail in the supplement [7].

5 Analysis

The domain of strategic management was represented in StrategyAdvisor Cloud using the different visual languages of the applied methodology and eventually turned

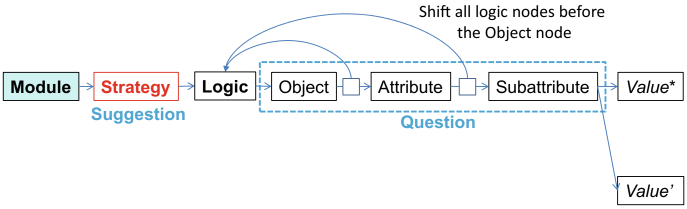


Fig. 6 Node types (vocabulary) and the sequence of nodes (grammar) in the mind map (Stage 1)

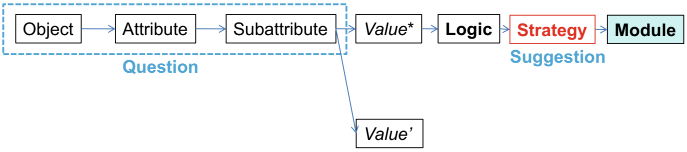


Fig. 7 Node types (vocabulary) and the sequence of nodes (grammar) in the domain objects map (Stage 2, DOM)

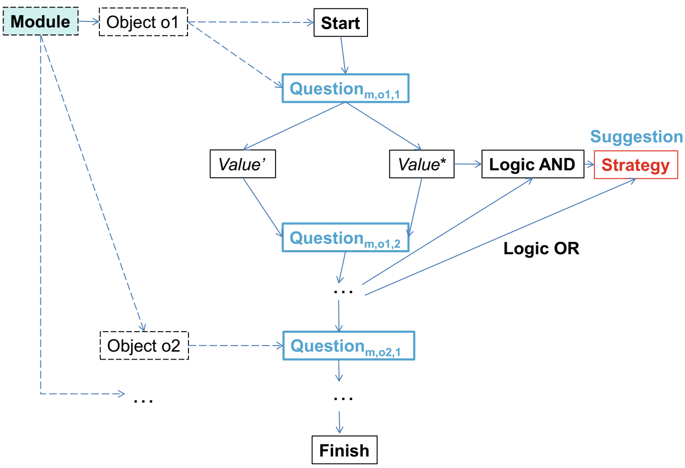


Fig. 8 Node types (vocabulary) and the sequence of nodes (grammar) in the rule map (Stage 3)

into a Web application designed, developed, and deployed, to serve as a digital consultant for strategic management. This section analyzes the expert system and the process through which it was developed.

The developed StrategyAdvisor Cloud expert system (Figs. 1, 2, and 3) suggests strategies for middle and top managers of companies to increase their profits, based on the facts that they provide. For various categories and functions of business planning (such as value chain, channel, and customer) (Fig. 1), the system gathers facts through convenient questions (Fig. 2). Once the system obtains sufficient facts to reach conclusions, it displays the suggested profit patterns as actions to take (Fig. 3).

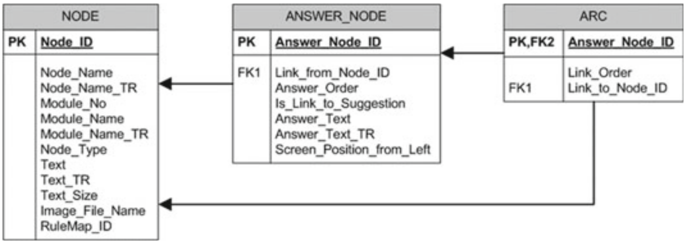


Fig. 9 Database structure used in the case study (Stage 4)

The methodology was implemented manually (without automation) in the case study, including the drawing of the graphs, transforming the graphs in the forward direction, creating the rule base in a database format, and constructing the text for questions and answers.

In populating the knowledge base of StrategyAdvisor Cloud, the *profit patterns* book, by Slywotzky et al. [14] at Mercer Management Consulting, was used as the principal source. The *profit patterns* book was selected as the pilot knowledge base for this study, primarily because the book is structured after patterns of profit, readily discussing the strategy rules, and presenting the strategy suggestions corresponding to each pattern. The principal challenge in transforming the book's knowledge into the rule base was making the transition from an essay style to a rule style, and this challenge was conveniently overcome through the mind maps of Stage 1.

The book *profit patterns* presents to business professionals 31 patterns observed to change the landscape of almost every industry. After the identification of key concepts and ideas in the book, strategy rules were extracted through mind maps, following the structure in (Fig. 6). The knowledge base was, therefore, initially represented in mind maps, then in domain object maps (DOM) (Fig. 7) and finally in rule maps (RM) (Fig. 8). A cloud software application, StrategyAdvisor Cloud (Figs. 1, 2, and 3), was then designed and created using the database structure in Fig. 9.

StrategyAdvisor Cloud is illustrated with examples from the **knowledge to product** profit pattern, where knowledge is converted into a product. This pattern is suitable for communicating the application of the methodology in the case study, because the case study itself actually follows this pattern: the knowledge of business thinker Adrian J. Slywotzky on strategic management [14] and the research team's knowledge and experience of expert system development and other fields (graph theory, information visualization, knowledge representation, database systems, strategic management) were transformed into the final product, the StrategyAdvisor Cloud software.

The final stage of the methodology is the creation of an expert system using a mainstream technology stack, including a programming language and its libraries. This stage is required if one does not wish to use an expert system language or engine as the production software. This last stage was pursued in the case study, and a cloud application was designed and created. The database structure for the rule base is shown in Fig. 9, the home modules page is given in Fig. 1, and sample snapshots for fact gathering and strategy suggestions are illustrated in Figs. 2 and 3, respectively. The developed system mimics a consultancy service for strategic management and can be used by any company anywhere in the world, free of charge. Technology selection decisions for the cloud software are explained in detail in the supplement [7].

The StrategyAdvisor Cloud software reads the strategy rules from the rule base that are stored in a relational database. When the software is initially launched, it displays the module selection page, where the user is expected to select one module at a time to run. The modules in the system are *channel, customer, value chain, knowledge, mega, organization*, and *product*.

The facts are gathered in the StrategyAdvisor Cloud through the *question pages*. Figure 2 shows a question page during the execution of the **knowledge** module.

For each question in the StrategyAdvisor Cloud, the answers are rated on a scale from 0 to 10. The boundary values of 0 and 10 are labeled with descriptive text, such as **Limited** and **Many** in Fig. 2.

Once all the relevant questions in a module have been completed, computer reasoning is carried out, and the suggestion window is displayed (Fig. 3). The window successively lists every suggestion that was found "applicable" based on the computation of a ratio for that suggestion. The numerator of this ratio is the number of arcs to that suggestion in the rule map traversed (an answer arc is traversed if the answer to its question takes values 1, 2, 3, or 4). The denominator of the ratio is the number of arcs in the rule map that terminate at that suggestion. Instead of displaying the scores, a threshold score can also be used. For example, if the score of a suggestion is greater than or equal to the crucial value of 0.6 (60%), then the suggestion can be displayed in the suggestion window. This particular threshold value is also used by Balch et al. [2] and de Souza et al. [15].

The suggestion window presents the strategies suggested to the user by StrategyAdvisor Cloud. For example, in Fig. 3, the **Knowledge to product** strategy is suggested, and the actions to take for the strategy are outlined.

The supplement [7] presents an assessment of StrategyAdvisor Cloud by the research team. The assessment provides strong evidence for the potential success of StrategyAdvisor Cloud, because all but one of the applicable criteria listed by Nurminen et al. [11] are satisfied by StrategyAdvisor Cloud.

6 Conclusions

A multistage methodology, built on knowledge graphs, was applied solve the challenges of knowledge acquisition in developing SMES. The applied methodology caters to the priorities and tasks of each team member in an expert systems project. Graph transformation algorithms allow the representation of the same rule base in various graph structures, allowing flexibility in the rule base development process. The methodology was then employed in a case study in which an expert system was developed to support strategic decision-making. The project experience and a formal assessment against success criteria suggest that the methodology has enabled the rapid development and deployment of an usable and potentially successful expert system. The special importance of StrategyAdvisor Cloud in small-medium enterprises (SMEs) is discussed in the supplement [7].

7 Future Work

The present study can be extended to the future with respect to both methodology and application, as summarized below and elaborated in the supplement [7].

– Visual specification of more complicated rules in knowledge graphs.
– Use of fuzzy reasoning and multi-criteria decision-making (MCDM) for scoring and ranking suggestions.
– Automated extraction of information and representation in knowledge graphs using text mining techniques.
– Integration of all the stages of the methodology in a single expert system modeling software.
– Automatically generate the source code of the desktop or Web applications and the executable of the desktop application.
– Analysis of logged data that users input into the system.
– Extension of the knowledge base to include other information sources.
– Addition of new modules in other functions of management, such as finance, supply chain management (SCM), and human resource management (HRM).
– Usability tests, which would include user surveys, to assess the applicability of StrategyAdvisor Cloud in business and industry.

Data Availability The domain objects map (DOM) knowledge representation of the rules is publicly available under https://ertekprojects.com/new-knowledge-in-strategic-management/data-yed-graphs/ as yEd graph files.

Acknowledgements This research was based on Startup Grant Funding 31B127 of the United Arab Emirates University. The authors thank Ashraf Khalil, Matloub Hussain, Aysha Al-Kaabi, Damla Uygur, Ceylin Özcan, Özge Onur, Ceren Atay and Gizem Kökten, Gül Tokdemir, and Nihat Kasap for their suggestions to improve the paper and help with the literature search, Cem Kanpara for redrawing the DOM in yEd software, Kamil Çöllü and Soner Ulun for their support in editing, Richard Wilkinson for his help in proofreading, and Clive Spenser from Logic Programming Associates (creator and distributor of VisiRule and Win-Prolog software products) for his extensive help and support throughout the project.

References

1. Amailef K, Lu J (2013) Ontology-supported case-based reasoning approach for intelligent m-government emergency response services. Decis Supp Syst 55(1):79–97
2. Balch RS, Schrader SM, Ruan T (2007) Collection, storage and application of human knowledge in expert system development. Expert Syst 24(5):346–355

3. Brenas JH et al (2018) Applied graph transformation and verification with use cases in malaria surveillance. IEEE Access 6:64728–64741. https://doi.org/10.1109/ACCESS.2018.2878311
4. Corral José Maria Rodriguez et al (2019) A study on the suitability of visual languages for non-expert robot programmers. IEEE Access 7:17535–17550. https://doi.org/10.1109/ACCESS.2019.2895913
5. Gupta I, Nagpal G (2020) Art Intell Expert Syst. Stylus Publishing, LLC, Sterling, VA
6. Irdesel I (2008) Strategy advisor: an expert system for strategic management consulting. MA thesis. Istanbul, Turkey: Graduate School of Engineering and Natural Sciences, Sabanci University
7. Irdesel I et al (2022) Supplement for "development of a web-based strategic management expert system using knowledge graphs. https://ertekprojects.com/ftp/supp/16.pdf
8. Kernbach S, Eppler MJ, Bresciani S (2015) The use of visualization in the communication of business strategies: an experimental evaluation. Int J Business Commun 52(2):164–187
9. McNulty P, Chembrakalathil V (2019) Overview of business rules management technologies at SAP. Tech Rep SAP
10. Nissen ME (2019) Initiating a system for visualizing and measuring dynamic knowledge. Technol Forecast Soc Change 140:169–181
11. Nurminen JK, Karonen O, Hätönen K (2003) What makes expert systems survive over 10 years—empirical evaluation of several engineering applications. Expert Syst Appl 24(2):199–211
12. Oracle (2020) Oracle business process management 12.2.1. 2020
13. Qin Y et al (2018) Towards an ontology-supported case-based reasoning approach for computer-aided tolerance specification. Knowl Based Syst 141:129–147
14. Slywotzky A et al (1999) Profit patterns: 30 ways to anticipate and profit from strategic forces reshaping your business. Random House, New York
15. de Souza HJC et al (2012) Project management maturity: an analysis with fuzzy expert systems. Brazil J Operat Product Manage 9(1):29–41
16. Surma J (2015) Case-based approach for supporting strategy decision making. Expert Syst 32(4):546–554
17. Tan CF et al (2016) The application of expert system: a review of research and applications. ARPN J Eng Appl Sci 11(4):2448–2453
18. Wagner WP, Najdawi MK, Chung QB (2001) Selection of knowledge acquisition techniques based upon the problem domain characteristics of production and operations management expert systems. Expert Syst 18(2):76–87
19. Zabukovec A, Jaklic J (2015) The impact of information visualisation on the quality of information in business decision-making. Int J Technol Human Interact 11(2):61–79

Received Power Analysis In Non-interfering Intelligent Reflective Surface Environments

Khalid Sheikhidris Mohamed, Mohamad Yusoff Alias, Mohammed E. A. Kanona, Mohamed Khalafalla Hassan, and Mutaz Hamad Hussein

Abstract The use of small cells, millimetre waves (mmWaves), and ultra-massive multiple-input-multiple-output (MIMO) systems have reshaped the future of wireless communication and showcased fifth generation (5G) to become the most promising communication system. Despite the significant quality of service (QoS) enhancements these technologies brought, the propagation channel challenge remains unsolvable. This is because equipments bear zero knowledge on both channel status and its effects on the propagating signals. It is therefore evident that these key technologies alone would not be sufficient to enable intelligent wireless platforms for the next generation communication, i.e. holographic communication, everything-to-everything (E2E). Intelligent reflective surfaces (IRSs) are thin and low-cost sheets yet very effective elements consisting of mini passive elements that can either manually be programmed or independently using artificial intelligence (AI) to alter the phase shifts of the impinging signals and proactively control the propagation channel. This paper investigates the impacts of using multiple IRS modules on the received power in an non-line-of-sight outdoor wireless environment considering. The experimental results show that received power can be enhanced by 9%.

Keywords Intelligent reflective surface · 6G · Reconfigurable metallic surface · Fading

K. S. Mohamed · M. E. A. Kanona · M. K. Hassan · M. H. Hussein
Innovation research and development center (IRDC), The Future University, Khartoum, Sudan
e-mail: khalidkaradh@fu.edu.sd

M. E. A. Kanona
e-mail: m.kanona@fu.edu.sd

M. K. Hassan
e-mail: memo1023@hotmail.com

M. H. Hussein
e-mail: mutaz_hamed@fu.edu.sd

M. Y. Alias (✉)
Multimedia University, Cyberjaya, Malaysia
e-mail: yusoff@mmu.edu.my

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_49

1 Introduction

Wireless connectivity now a days is easily accessible because of the massively large footprints of cellular and WiFi networks. This increases the human work productivity, centralizes the social interchange, and facilitates the data exchange in other sectors, e.g. business, health care, technology, etc. With regards to that, a 100 Km^2 city for instance, requires about 3 cellular towers (depending on the transmit power) to provide coverage for the entire city. A typical cell phone that located in about 10 metres from the base station may receive about 0.000001% of typical frequency transmit power in non-line-of-sight (NLOS) environment because of its receiving antenna size, and the propagation constraints. The remaining power/energy either fades in the channel or gets received by other terminals, i.e. interferers. The fifth generation (5G) systems rely on massive multiple-input multiple-output (MIMO) technology to provide increased gain signals and higher data rates to users (e.g. 20 Gbps in the downlink as reported by [1]).

To date, fifth generation (5G) system's footprint is relatively small because it has gradually been rolling out since 2019, yet the speculations of the sixth generation (6G) communication system have already begun in-order to look for ultra-fast speeds and enormous capacity advancements over the 5G. Subsequently, several initiatives are currently speculating the 6G communication system to describe features beyond the capability of the existing 5G communication system. This is because the data traffic has increased over 20 time than of 10 years before, and it is expected to reach up to 5 zettabytes (ZB) by 2030 [2]. In comparison to the fourth generation (4G) communication system, 5G offers 1000 times increased capacity, and about 10 Gbps in uplink and about 20 Gbps in the downlink as reported by [1]. However, applications such as holographic projection [3], remote surgery, cell-free systems [4], vehicle-to-everything (V2X) [5], high definition 3D maps indicate that massive amount of data will be shared in the next 10 years, which cripples the 5G from embracing such demands. Therefore, 6G will be called upon to make for the upcoming step.

The eruption of data exchange as discussed above will eventually lead to massive deployment of short range radio networks (e.g. leq 10 m), thus, hyper-dense networks and more energy consumption and pollution. 5G base stations that support millimetre waves (mmWaves) for instance consume nearly four-folds of 4G base stations (18.9 kW peak consumption compared to 7.3 kW). Such trend is expected to continue in 6G leading to more energy pollution, increased energy consumption, and huge operational costs. The energy consumption/efficiency therefore shapes up to be a great challenge in future radio communication. This can be reduced by either decreasing the cell size and optimizing the transmission power, i.e. small cells, and/or forcing small cell base stations (BSs) into sleep modes [6], while the neighbouring BSs slightly increase their power to serve the users within the sleep zone. The selection of the BSs usually depends on two policies, (1) random selection where all BSs are on standby mode and might sleep anytime, (2) strategic selection where the sleep probability depends on the traffic load of that BS. The reduced energy depends on four modes (on, standby, sleep, and off), each has a different wake up time and power consumption. The BS lowers its power expenditure by using linear processing techniques (e.g. maximal ratio transmission, maximal ratio combining, etc.) [7] while EE is enhanced by increasing the directivity of the signals, thus, higher spectral efficiency (SE). Although this might be applicable for 5G systems but it may not be achieved for next generation networks (NGNs) because the energy consumption will surely increase. Thus, a trade-off between the energy consumption, coverage, and the quality of service (QoS) exists for the NGNs.

Intelligent reflective surfaces (IRSs) on the other hands, emerged as a promising paradigm that can solve for the propagation shortcomings in NGN radio communication [8]. These metasurfaces comprise of mini low-cost passive elements, printed dipoles, and phase shifters used to intelligently alter the phase shifts of the impinging signals in order to proactively control the propagation environment [9]. IRSs are programmed using external interfaces (controllers) that bear some knowledge on the characteristics of the impinging waves (e.g. polarization, phase, frequency, and amplitude) [10]. Nevertheless, IRSs should have few characteristics: (1) automatically control their response to the impinging signals, (2) dynamically adapt to different wave behaviours (e.g. refraction, reflection, etc.). Also, IRSs do not need external power supply, i.e. reduced energy consumption and are considerably cheap in comparison to conventional cell splitting and sectorization methods. Using conventional signal processing and artificial intelligence (AI), the IRSs can be autonomously controlled to achieve optimized radio propagation hence, improved signal reception.

In this paper, the received power is analysed considering the use of four IRSs deployed at different locations in the propagation environments and 5 users located randomly. The simulation adopts the 3rd generation partnership project (3GPP)-like channel model presented in [11] and considers an outdoor propagation environment. The findings on this paper represent a foundation for enhanced signal reception in beyond 5G networks incorporating IRSs to embrace the new desires at both societal and individual levels.

The rest of the paper is organized as follows: Sect. 2 presents a discussion about the NGNs, Sect. 3 discusses the simulation modelling, Sect. 4 presents the findings of this paper, and Sect. 5 concludes the paper.

2 Next Generation Networks

Considering the previous discussion, 6G will profoundly embrace more degree of freedom in wireless communication, enhance the antennas' abilities to collect more wireless signals, and deliver an unprecedented network capacity. It will also enhance the security, and might make much more impact in wireless power transfer and energy harvesting areas to enable phones to automatically charge themselves using the radio waves and laser beams [12]. In about 10–15 years from now, super-smart cities, Internet of everything (IoE) [13] where enormous amount of data and information and autonomous services are available for mobile phones, flying vehicular technology [14], etc. will all be available because of 6G systems. This is motivated by the development of industry X.0 concept [2] which is comprised of millions of robots and massive ultra reliable low latency communication (URLLC) services such as smart health care with ultra high security, ultra holographic conferencing, and wireless 2.0 [15] and intelligent radio environments [16] where networks and environments are both customized. Subsequently, due to the enormous complexity in signals processing, resource allocation, computational capability, identifying and dealing with interference patterns and dynamics, etc., artificial intelligence (AI) and machine learning (ML) [17] have to support nearly every corner in 6G [15].

Similar to all other systems, 6G will need some efforts before any phase of implementation is started, some topics have already been addressed in 5G and some are still open such as the backhaul subsystem that is supposed to handle the unprecedented data traffic, the capability of the connected devices in-terms of computational performance and power consumption, the dynamically changing topologies such as V2X, the operating frequency bands and its' modelling in-terms of interference dynamics and signals susceptibility to penetration losses, resource distribution or sometimes referred to as resource as a service (RaaS) [2] to give rise to the concepts of physical and virtual network integration, i.e. network slicing, and extremely short range communication otherwise known as whisper radio [18].

2.1 Favourable Propagation Environments

The limited knowledge and/or imperfect channel state information as well as the limitation of the contemporary approaches make the IRS an appealing solution to bring significant performance enhancements, and at the same time combat the propagation implications. When embedding IRSs into the existing communication environment (indoor and/or outdoor), the signal reception, maximum sum rate, and spectral efficiency are significantly enhanced because of the establishment of the favourable radio propagation and more controllable and optimized propagation. It is understood that the radio waves impinging on large surfaces can coherently add-up and be reshaped at any point on the contiguous surface to combat the impact of multipath fading as illustrated in Figure 1. Noise from the surrounding environment does not affect IRSs, and neither analogue/digital converters nor amplifiers are needed. Consequently, an IRS neither amplifies nor produces noise while intercepting the reflecting signals and offers completely independent duplex transmission. Particularly, IRSs have maximum band response, hence, they can theoretically function at any operating frequency. Up to the authors knowledge, the effectiveness of using the IRSs with traditional signal processing approaches is still theoretical hence, this paper opens the doors for enabling favourable propagation in high communication demand areas, i.e. NGNs.

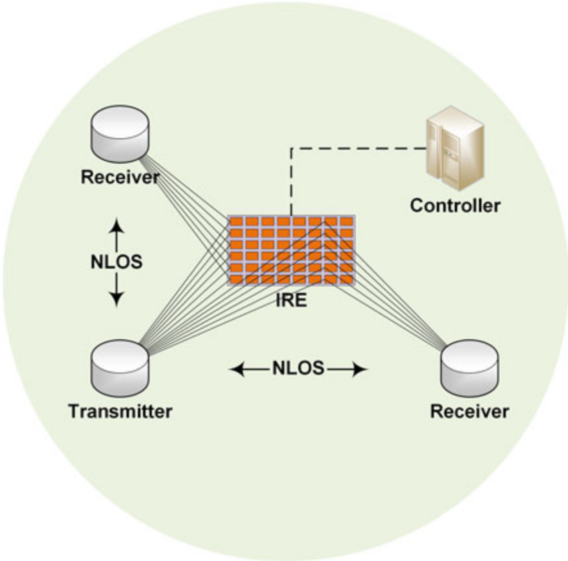


Fig. 1 IRS module illustration

2.2 Related Works

Many next generation technologies have been paid for the development of wireless networks as a result of the proliferation of smart devices and emerging applications [19]. Despite the fact that commercial 5G has only recently become widely available in some countries, there have been preliminary efforts from academia and industry to develop 6G systems. A large number of devices and applications emerge in such a network, as well as heterogeneity of technologies, architectures, mobile data, and so on, and optimizing such a network is critical. However, the main challenges of the next generation networks are: quality of service; management; security; economic; and transition, is the need to ensure a smooth transition to the new a data network.

IRS is a novel reflective radio technology that is attracting growing attention in recent years. The reflective array concept was first proposed in [20] then, it has been introduced in the wireless communications research community [21]. Form the physical view the IRS envisioned as a large planar array of passive reflecting antennas with unique structure to achieve different communication goals [22]. In addition, software defined networks and other artificial intelligence mechanism were used to change and control the electromagnetic properties of each scattering element. In the near future of wireless networks, this smart integration of radio environment and network optimization paradigm is expected to play a big role [23]. Reflective devices that do not use expensive and power-hungry active components have become popular for transmitting signals to their receivers or improving the transmission of primary communication systems using EM scattering of radio waves. The most significant benefits of IRS in wireless communications are flexibility, programmability, sustainability, ease of deployment, and capacity and performance enhancement.

Most studies [24, 25] have considered the change as a phase shift only to the incident signal, resulting in an IRS that consumes no transmit power. When direct communications have poor quality, an IRS intelligently configures the wireless environment to help the transmissions between the sender and receiver. IRSs can be placed on walls, building facades, and ceilings, as in [21]. Under proportional rate constraints, the spectral efficiency of an IRS-aided multi-user system was investigated, and an iterative optimization framework was proposed in [26]. This optimized the transmit covariance and IRS phase shifts to maximize the secrecy rate for an IRS-assisted multi-antenna systems.

IRSs have also been used in mmWave communication systems, where a BS with a few active antennas illuminates a large IRS nearby [27]. Massive MIMO beamforming gains can be achieved by increasing the number of passive elements at the IRS without increasing the number of active antennas at the BS [28]. However, by placing the IRS very close to the BS, these works assume a lossless fixed connection between the BS and the IRS. The problem of joint active and passive precoding design for IRS-assisted mmWave systems, where multiple IRSs are deployed to assist data transmission from the BS to a single antenna user, was studied by the author in [29]. The evolution of reflective arrays to the IRS as well as the communication model of an IRS-assisted multi-user MISO system and how it differs from traditional multi-antenna communication models has been discussed. An MMSE-based channel estimation protocol at a 2.5 GHz operating frequency was proposed to estimate the IRS assisted links. The findings revealed that IRSs can assist in the creation of effective virtual LOS paths to improve the robustness of mmWave systems in the face of blockages.

The authors of [30], investigate the uplink outage performance of IRS-assisted nonorthogonal multiple access (NOMA) by considering the general case where all users have both direct and reflection links and all links fade to Nakagami-m. In another scenario, authors in [31] investigate the spectral and energy efficiency [32] of an intelligent reflecting surface (IRS)-assisted multiple-input single-output (MISO) downlink system with hardware flaws. The findings show that the performance of both the AP and the IRS is hampered by an increasing number of elements.

The authors of [33] present instructions for embedding arrays of low-cost antennas into a building's walls in-order to passively reflect incident wireless signals. Three parabolic antennas loaded with single-pole four-throw switches are used in the prototype, which can achieve 64 different reflection configurations. Author in [34] shows how to make a reflect-array with 224 reflecting units by loading a microstrip patch element with an electronically controlled relay switch. Variable capacitors are integrated into the reflector panel to continuously tune the phase responses of the reflecting units.

3 Modelling

The simulation presented in this paper considers the use of several IRS modules whereby their locations are determined by the coverage area radius. The distance in between the an IRS and the other is dependant on the number of IRSs, i.e. the more IRSs is the smaller the distance in between. The random location and separation distance selection pave the way to later on understand the effects on the QoS and help in selecting the optimal location that serves the QoS in the best way. An essential factor in this simulation is the channel model whereby newer channel models should be considered in NGN simulations. This is because these networks are required to support 5G operation within frequency band ranges of up to 100 GHz should. Although 3 dimension (3D) models give more realistic results, this paper considers a 2D 6GHz propagation. In that regard, the propagation loss is obtained by the following [35]:

$$Pl(d)[\mathrm{dB}] = 20log_{10}(\frac{4\pi f d_0}{c})[\mathrm{dB}] + 10nlog_{10}(\frac{d}{d_0}) + X_\sigma \tag{1}$$

where d_0 is the free space reference distance, d is the distance between transmitter and receiver, n denotes the path loss exponent, f is the frequency in GHz, c is the speed of light, and X_σ is the zero-mean Gaussian random variable with a standard deviation σ in dB (shadowing effect).

The received power can therefore be analysed in a non interfering environment, i.e. interference is not considered. However, this is subjected to the channel state.

3.1 Simulation Setup

The simulation in this paper considers several factors that are in-lined with the 3GPP standards. The simulations considers four IRSs are distributed in the area and 5 random users are static to one location per iteration whereby each user is assigned to a single IRS module. The rest of the simulation parameters are shown in Table 1.

4 Experimental Results

The results shown in Figure shows a comparison of the pathloss of signals received from the base station through the IRS modules and without them. Looking at user 1 for instance, it is easily understood that the pathloss of at least one IRS is less than of receiving from the BS directly (Fig. 2).

This is later reflected at the received power analysis in Figure 3 where the users receive the power without using the IRS modules. The figure shows that the maximum

Table 1 Simulation parameters

Parameter	Value
Frequency	6 GHz
Number of IRSs	4
IRSs locations	Random
Number of users	5
Users locations	Random
User association	Single IRS per iteration
Iterations	100

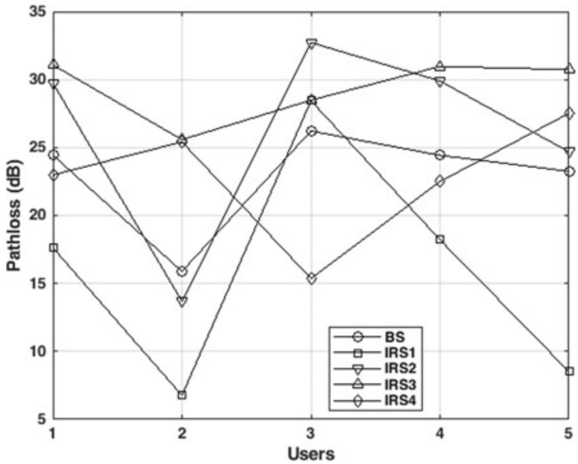


Fig. 2 Pathloss analysis

received power is received by user 5 which is about 0.017 dB and the minimum is received by user 2 which is about 0.001 dB.

Comparing these results to Fig. 4 where the received power when using the IRS is presented, it can be viewed that the received power is significantly enhanced. The same users highlighted in Fig. 3 now receive improved signals when using IRS modules by approximately 9% which is an impressive amount considering these simple simulations setups

5 Conclusion

Intelligent communication environments generally contain three major elements: (1) IRS embedded objects (the channel), (2) computational learning algorithms, and (3) the main network elements (base stations and user terminals). Although IRS-aided communication is the paradigm shifting notion, it is ultimately challenging to

Fig. 3 The average received power without using IRS modules

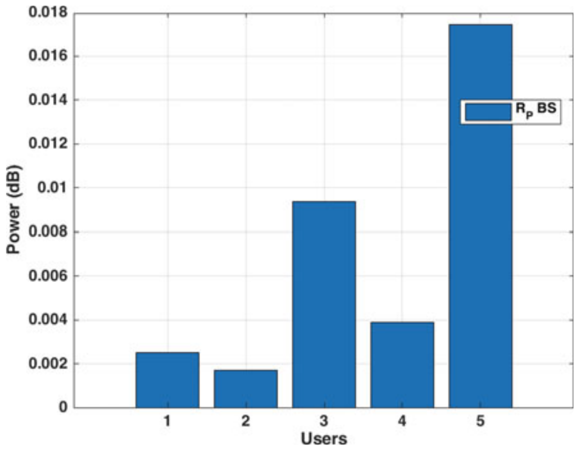
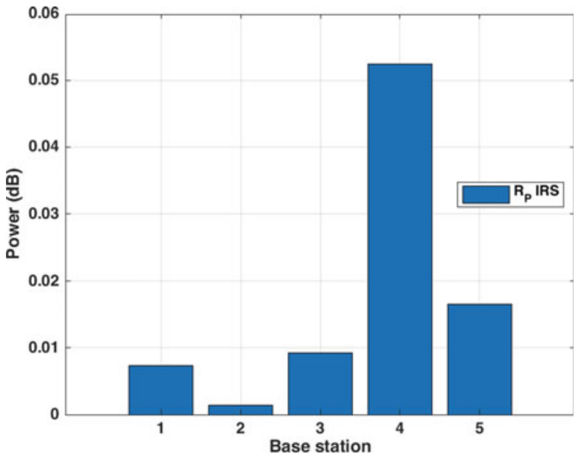


Fig. 4 The average received power through IRS modules



design and build a technology that can autonomously adapt to the environment scale. Coming up with a factual IRS that can simply work on its own once embedded in the environment is considerably challenging. At a certain level, ML approaches have to be used to exploit the existing knowledge, analyse signals behaviour and channel response, and support the decision making in order to assist IRSs to adapt to any environment and execute intelligent tasks without being controlled or reprogrammed. Although the analysis on this paper might not be sufficient to conclude the suitability of the IRSs, but to some extent it shows that IRSs can be effective options for free space propagation environments whereby it contribute in enhancing the received power by approximately 9%. The future scope of this work includes exploring the performance in dense areas and/or with increased number of IRSs.

Acknowledgements This research is a joint effort between Future University and Multimedia University. The authors are grateful to all those who contributed to the study by sharing their experience and knowledge. We would like to acknowledge Telekom Research and Development Sdn. Bhd. (TM R&D) for providing financial sponsorship to facilitate this research project under TM R&D Research Grant 2021 (Project Code: MMUE/220001).

References

1. GT 38.913 (2017) Technical specification group radio access network; study on scenarios and requirements for next generation access technologies
2. Tariq F, Khandaker M, Wong KK, Imran M, Bennis M, Debbah M (2019) A speculative study on 6g. arXiv preprint arXiv:1902.06700
3. Daniel IS (2017) Apparatus, system and method for holographic video conferencing. US Patent 9,661,272
4. Ngo HQ, Ashikhmin A, Yang H, Larsson EG, Marzetta TL (2017) Cell-free massive MIMO versus small cells. IEEE Trans Wireless Commun 16(3):1834–1850
5. Chen S, Hu J, Shi Y, Peng Y, Fang J, Zhao R, Zhao L (2017) Vehicle-to-everything (v2x) services supported by LTE-based systems and 5g. IEEE Commun Stand Magaz 1(2):70–76
6. Liu C, Natarajan B, Xia H (2015) Small cell base station sleep strategies for energy efficiency. IEEE Trans Vehi Technol 65(3):1652–1661
7. Prasad KSV, Hossain E, Bhargava VK (2017) Energy efficiency in massive MIMO-based 5g networks: opportunities and challenges. IEEE Wireless Commun 24(3):86–94
8. Wu Q, Zhang R (2019) Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. IEEE Trans Wireless Commun 18(11):5394–5409
9. Xu D, Yu X, Sun Y, Ng DWK, Schober R (2020) Resource allocation for IRS-assisted full-duplex cognitive radio systems. IEEE Trans Commun 68(12):7376–7394
10. Bariah L, Mohjazi L, Muhaidat S, Sofotasios PC, Kurt GK, Yanikomeroglu H, Dobre OA (2020) A prospective look: key enabling technologies, applications and open research topics in 6g networks. IEEE Access 8:174–792, 174–820
11. Haneda K, Zhang J, Tan L, Liu G, Zheng Y, Asplund H, Li J, Wang Y, Steer D, Li C (2016) 5g 3g pp-like channel models for outdoor urban microcellular and macrocellular environments. In: IEEE 83rd vehicular technology conference (VTC spring). IEEE, pp 1–7
12. David K, Berndt H (2018) 6g vision and requirements: is there any need for beyond 5g? IEEE Vehi Technol Magaz 13(3):72–80
13. Jara AJ, Ladid L, Gómez-Skarmeta AF (2013) The internet of everything through ipv6: an analysis of challenges, solutions and opportunities. JoWua 4(3):97–118
14. Steder B, Grisetti G, Stachniss C, Burgard W (2008) Visual slam for flying vehicles. IEEE Trans Robot 24(5):1088–1093
15. Gacanin H, Di Renzo M (2020) Wireless 2.0: towards an intelligent radio environment empowered by reconfigurable meta-surfaces and artificial intelligence. arXiv preprint arXiv:2002.11040
16. Di Renzo M, Debbah M, Phan-Huy D-T, Zappone A, Alouini M-S, Yuen C, Sciancalepore V, Alexandropoulos GC, Hoydis J, Gacanin H et al (2019) Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come. EURASIP J Wireless Commun Netw 2019(1):1–20
17. Khairi MHH, Ariffin SHS, Latiff NMA, Yusof KM, Hassan MK, Al-Dhief FT, Hamdan M, Khan S, Hamzah M (2021) Detection and classification of conflict flows in SDN using machine learning algorithms. IEEE Access 9:76, 024–037
18. Xing Y, Rappaport TS (2018) Propagation measurement system and approach at 140 ghz-moving to 6g and above 100 ghz. In: IEEE global communications conference (GLOBECOM). IEEE pp 1–6

19. Pham QV, Nguyen DC, Mirjalili S, Hoang DT, Nguyen DN, Pathirana PN, Hwang WJ (2020) Swarm intelligence for next-generation wireless networks: recent advances and applications. arXiv preprint arXiv:2007.15221
20. Berry D, Malech R, Kennedy W (1963) The reflectarray antenna. IEEE Trans Antennas Propagat 11(6):645–651
21. Wu Q, Zhang R (2019) Towards smart and reconfigurable environment: intelligent reflecting surface aided wireless network. IEEE Commun Magaz 58(1):106–112
22. Liang Y-C, Long R, Zhang Q, Chen J, Cheng HV, Guo H (2019) Large intelligent surface/antennas (LISA): making reflective radios smart. J Commun Inform Netw 4(2):40–50
23. Zappone A, Di Renzo M, Debbah M (2019) Wireless networks design in the era of deep learning: model-based, AI-based, or both? IEEE Trans Commun 67(10):7331–7376
24. Basar E, Di Renzo M, De Rosny J, Debbah M, Alouini MS, Zhang R (2019) Wireless communications through reconfigurable intelligent surfaces. IEEE Access 7:116, 753–773
25. Jung M, Saad W, Kong G (2021) Performance analysis of active large intelligent surfaces (LISS): uplink spectral efficiency and pilot training. IEEE Trans Commun 69(5):3379–3394
26. Shen H, Xu W, Gong S, He Z, Zhao C (2019) Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications. IEEE Commun Lett 23(9):1488–1492
27. Jamali V, Tulino AM, Fischer G, Müller R, Schober R (2019) Intelligent reflecting and transmitting surface aided millimeter wave massive MIMO. arXiv preprint arXiv:1902.07670
28. Elamin NIM, Abd Rahman T (2015) 2-element slot meander patch antenna system for LTE-WLAN customer premise equipment. In: 2015 IEEE-APS topical conference on antennas and propagation in wireless communications (APWC). IEEE, pp 993–996
29. Wang P, Fang J, Yuan X, Chen Z, Li H (2020) Intelligent reflecting surface-assisted millimeter wave communications: joint active and passive precoding design. IEEE Trans Vehi Technol 69(12):14, 960–973
30. Tahir B, Schwarz S, Rupp M (2020) Analysis of uplink IRS-assisted NOMA under nakagami-m fading via moments matching. IEEE Wireless Commun Lett 10(3):624–628
31. Zhou S, Xu W, Wang K, Di Renzo M, Alouini M-S (2020) Spectral and energy efficiency of IRS-assisted MISO communication with hardware impairments. IEEE Wireless Commun Lett 9(9):1366–1369
32. Almula HAF, Hamza ME, Kanona ME (2020) Improvement of energy consumption in cloud computing. In: 2020 International conference on computer, control, electrical, and electronics engineering (ICCCEEE). IEEE, pp 1–6
33. Welkie A, Shangguan L, Gummeson J, Hu W, Jamieson K (2017) Programmable radio environments for smart spaces. In: Proceedings of the 16th ACM workshop on hot topics in networks, pp 36–42
34. Tan X, Sun Z, Koutsonikolas D, Jornet JM (2018) Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays. In: IEEE INFOCOM 2018-IEEE conference on computer communications. IEEE, pp 270–278
35. Li SD, Liu YJ, Lin LK, Sheng Z, Sun XC, Chen ZP, Zhang XJ (2017) Channel measurements and modeling at 6 ghz in the tunnel environments for 5g wireless systems. Int J Antennas Propagat

Measuring Vital Signs for Virtual Reality Health Application

Leonel D. Deusdado, Rui P. Lopes, Alexandre F. J. Antunes, and Júlio C. Lopes

Abstract Smart devices are extremely useful nowadays and incorporate a variety of tools. Their advanced technology allows them to multitask on the same device, being because of this and due to the convenience, agility, and accuracy they bring, they are becoming increasingly popular. Taking advantage of this, it is possible to use this technology to assist in many areas and for a variety of objectives. In areas such as the health sector, new devices are increasingly being used to aid in the treatment of patients in a preventive way, such as with degenerative mental diseases. This research explores the use of a smartwatch to capture heart rate, store it in a database, and display it on a web page, all in real time. The future significance of such a development is based on its connection with another segment of research that attempts to employ virtual reality to aid in the treatment of schizophrenia with a serious game, being able to perceive schizophrenics' body behavior for health analysis.

Keywords Health · Vital sign · Measurement · Storage · Real time · Virtual reality

1 Introduction

People are increasingly measuring their vital signs, whether for exercise or body analysis for health monitoring. This is due to the increasing use of smart devices that can quickly provide such data and a solid assessment of what is being read as

L. D. Deusdado (✉) · R. P. Lopes · A. F. J. Antunes · J. C. Lopes
Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança (IPB), Campus de Santa Apolónia, 5300-253 Bragança, Portugal
e-mail: leodeus@ipb.pt

R. P. Lopes
e-mail: rlopes@ipb.pt

A. F. J. Antunes
e-mail: alexandrejarosz@ipb.pt

J. C. Lopes
e-mail: juliolopes@ipb.pt

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_50

well. The smartwatch is a popular technology these days. It is commonly used by athletes to conduct thorough measures of calorie consumption, blood oxygen level, and several others, including heart rate, which is the focus of this research.

It is also beneficial in a medical context and offers various advantages for ongoing and extensive analysis. It is usual for data to be stored and utilized for behavioral verification as technological gadgets become more integrated. The project's purpose is to better understand how schizophrenics act by measuring their heartbeats while using a virtual reality serious game application and with this also aid in their treatment. The next sections will go over such applications in further detail as well as present the use of a smartwatch for heartbeat rate measurement.

2 State of the Art

Digital technology is becoming more prevalent in everyday life and more accessible to everyone. Its evolution has been so rapid over the decades that humanity is constantly learning to reinvent it for new purposes. With devices that integrate multiple features, commonly referred to as smart devices, it is crucial to consider the quality and security they provide in order to make them usable. With this in mind, each branch in which they are used will necessitate various tests and analyzes to ensure their accuracy and effectiveness, whether for an external or internal body part.

It is essential for the medical sector to have ongoing scientific studies, such as the one conducted by Winiarski et al., who used inertial sensors attached to the body of an automotive multinational corporation worker to determine whether there is a relationship between ergonomic risk measures and health diseases via motion analysis in several routine tasks [1]. Also with the research of Alsaade et al., which used facial recognition and deep learning techniques in photos from social media to detect children with autism, in order to aid in their treatment, reintegrating them into society through early detection of the mental illness [2].

Measuring data for the inside of the human body is indispensable for detecting diseases and viruses, as well as assisting in the treatment of people who are already ill. A number of social changes and habits have resulted from the (COVID-19) pandemic. People were required to have body temperature check, either infrared temperature detectors or forehead thermometers [3], on most countries and most of commercial premises to ensure that the virus was not present. Rahaman et al. conducted a project estimating a person's calorie burn while doing certain types of activity without the use of a wearable device and then tested with results from those ones [4]. Basjaruddin et al., on the other hand, designed a device to measure the level of stress, based on the joint reading of vital signs such as body temperature, oxygen saturation, galvanic skin response, and heart rate, captured through sensors, as a way to solve the problem of low capture effectiveness of current medical devices that read each body information separately [5].

When it comes to schizophrenia treatment, which is the ultimate goal of the current project's development assistance, there is an increasing amount of technology-driven research. Abbas et al. conducted a research with schizophrenic and healthy individuals to assess motor functioning as a characteristic of schizophrenia over a two-week period, being trained by the study team to use a smartphone application that asked simple questions and captured a video of the participant's response, using the front-facing camera, making an analysis of their head movement by machine learning and computing vision [6].

It is then noticeable the increasing in the use of mobile devices to support medical treatment and public health, referred to nowadays as mHealth, but awareness is required not only for regular smart device users but also for doctors and other associated professions. Gupta et al. conducted a review with this approach, analyzing potential benefits of smartwatches in the clinical area, given their wide range of capturing metrics of the human body, whether with heartbeat rate, amount of oxygen, steps count, sleep tracking, (ECG or EKG), (HRR), or (HRV), realizing the great value and good results coming from the use of this technology [7]. The heart rate is one of the most frequently used features on smartwatches. This feature has a huge advantage in that it can be combined with many various purposes and provide support for other analyses. It also has the ability to tell how a person feels while performing an action. As in Nielsen's research, which uses a smartwatch to measure the heartbeat of children who have issues integrating and interpreting sensory information, that can lead to learning difficulties, then a haptic vest is used for sensory stimulation [8].

Another example would be in relation to (VR), such as the research of Hirzle et al. into this interaction, or the work of Salekin et al. on creating a tool for visualizing enormous and complex medical datasets, which uses the watch as an input method [9, 10]. Furthermore, Quintero et al. discussed the effectiveness of this use with physiological computing system for mental health therapies in their master thesis, analyzing (HRV) during slow-breathing relaxation exercises using a smartwatch and recognizing that medicine can benefit from mobile technology but emphasizing the need to be cautious of technical instabilities [11]. Another intriguing research was conducted by Dolu et al., which investigated how (VR) affects isometric muscle strength. They discovered that participants experienced less pain while using (VR) compared to the traditional exercise were capable of pushing themselves more and felt the time pass more quickly, and there was no difference in their (HRV), which is a metric measured using a smartwatch [12].

The current project uses a smartwatch to monitor the heart rate, as described in greater details in the following sections., with the intention of later integrating it with a (VR) application, sometimes in conjunction with the usage of a haptic vest, to aid in the treatment of schizophrenics [13–15].

3 Heart Rate

The capability to collect information directly from the user is essential for the operation of many smart device systems, especially when the application's objective is to monitor or examine bodily behavior metrics. Although they are not medical devices,

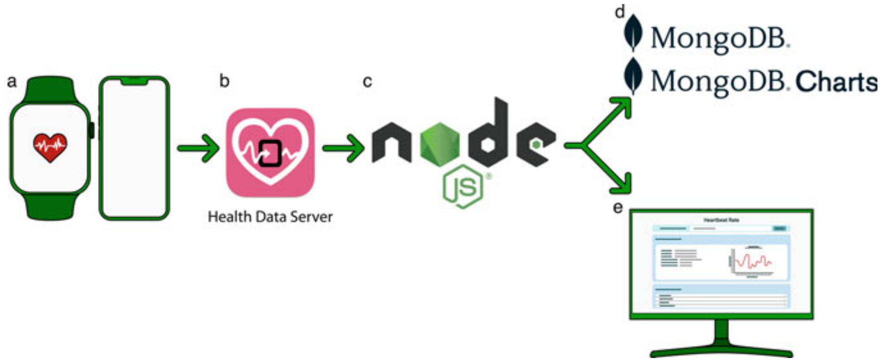


Fig. 1 Heart rate steps. **a** Smartwatch and smartphone. **b** Health data server. **c** Node.js. **d** MongoDB and MongoDB chart. **e** Webpage

smart devices are highly accurate nowadays and can provide a satisfactory estimate to the user, so there is no need to perform numerous medical tests in order to obtain basic information. Smartwatches, for example, that measure with good precision, are not overly intrusive, and do not require calibration.

With that in mind, the project's goal is to employ a smartwatch to capture metrics from user's heartbeat while they are using a (VR) application which focuses on supporting those who suffer from schizophrenia, a progressive mental disorder for which there is yet no cure [14, 16]. Such information will be recorded in a database for analysis of how people feel and when it happens, enabling analysts to determine in what way the (VR) application affects heart rate variations. Additionally, another branch of the study uses a haptic device to try to make the (VR) experience more immersive, which the data will also be used to analyze user's body feedback [13].

To accomplish this, it was essential to use and configure a smartwatch that could, as illustrated in Fig. 1, record heartbeats (connected with a smartphone) (Fig. 1a) and connect it to an application that could gather this information and send it to a desktop computer for which the Health Data Server Application [17] (Fig. 1b) was chosen. Those steps will be explained in Sect. 4.

The information was taken by coding with Node.js (Fig. 1c) [18], that will be discussed in Sect. 5, and then stored in a MongoDB (which provides MongoDB Charts) (Fig. 1d) [19] database, as it can be seen explained in Sect. 6, being all platforms used to generate the desired result. The last step was to display the data on a webpage (Fig. 1e), and it will be explained in the Sect. 7.

4 Heartbeat Rate Measurement

For the heartbeat rate measurement, it was chosen to use a smartwatch that could capture the heart rate quickly and reliably, and this is the first step of the process as it can be seen in Fig. 1a.

The Apple Watch Series 6 [20], from Apple Inc. [21], was selected as the smartwatch for the research because it includes advanced technology and a third-generation optical heart sensor in addition to an electrical heart sensor [22], besides being light and having a small area of contact with the user's body, it does not cause much discomfort; in other words, it is not quite intrusive. The watch contains an optical heart sensor for measuring heart rate, which can also use infrared light, green (LED) lights, and it has "built-in electrodes in the Digital Crown and the back of Apple Watch, which can measure the electrical signals across your heart" [23–25]. A disadvantage of adopting Apple's system is that in order for the watch to operate, you must always have the iPhone around, connected to the internet, and paired with the same Apple account.

The Health Data Server Application [17], created by Rexios, was used for heart rate capture because it effectively transmits the heartbeat rate fastly and is simple to use, and this is the second step of the process as it can be seen in Fig. 1b. It was installed on the iPhone and the watch itself. Figure 2 exhibits the application being used on the smartwatch while capturing the user's heartbeat (a), an user testing it with a haptic vest and our (VR) application displayed in Oculus Meta Quest 2 [26] (b). Part (c) of the image corresponds to what the user is experiencing in the Oculus, which is another branch of the project that has developed a (VR) serious game application [14] to aid in the treatment of schizophrenia, which will be tested with schizophrenics afterward. For such objective, the developer makes available the executable of various systems in a page on GitHub [27] that has instructions for using and downloading files [28].

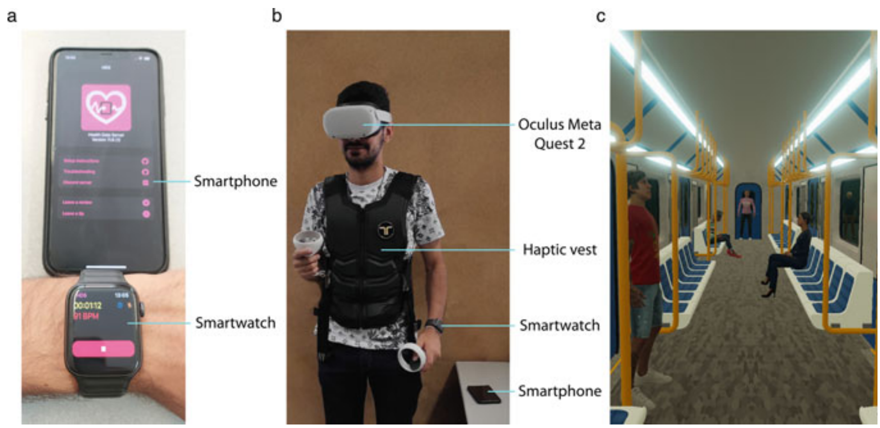


Fig. 2 Smartwatch usage process. **a** Smartwatch and smartphone with health data server. **b** Smartwatch, VR application and haptic vest. **c** VR application

5 Data Acquisition

The data acquisition from the smartwatch had to be received via an (IPv4) connection to the computer. In order to do this, programming code was required to obtain it and used to manage the entire data flow, starting with collecting it from the application until storing and displaying it, and this is the third step of the process as it can be seen in Fig. 1c.

For each user, it is started a new section on the website, followed by a new measurement within the smartwatch application, which starts recording each heartbeat. Our project's GitHub page hosts the open-source application code and also provides additional details on how to use and install the project [29]. The final step was to display the data on a webpage in a more simplified format. This page is being created locally on port 3000 on localhost. More information about the website will be provided in the Sect. 7.

6 Data Storage

It was necessary to use a database in a cloud for data storage that would allow for simplicity and agility, with the goal of transmitting data in real time, in addition to an online connection that would allow data transmission from a device. It was chosen a (NoSQL) oriented database which stores data in (JSON) [30, 31] format and was used to store user's heartbeat data from the smartwatch application, among other things, and this is the fourth step of the process as it can be seen in Fig. 1d.

Figure 3a exhibits the database for saving user data in MongoDB Atlas [32], MongoDB's cloud management. Several elements considered necessary to save were defined for this, such as the user's name, an identifier of the measurement section they are in, the start and final date and time of that section, and the section's duration in

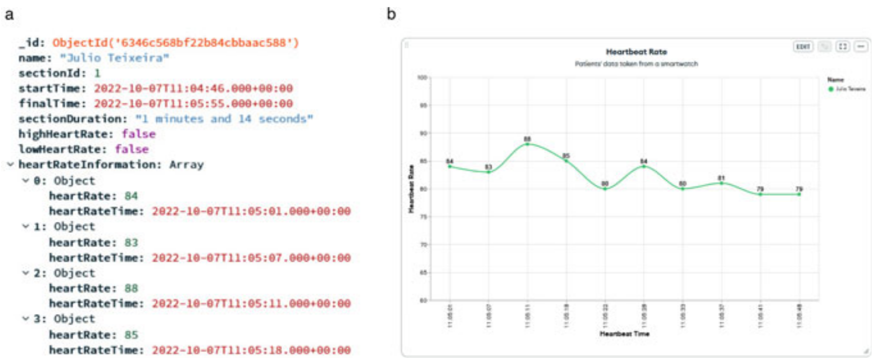


Fig. 3 Database configuration and display. **a** MongoDB storage. **b** MongoDB real-time chart

minutes and seconds. It also saves the possibility of the user surpassing the safe upper or lower limit of the heart rate, as specified by a doctor in a user analysis. Furthermore, the heartbeats are recorded along with the moment they were collected.

For Fig. 3b, a graph built on MongoDB Chart [33] is displayed with the same data as in Fig. 3a. Its configuration is determined by the heartbeat rate on the vertical Y axis and the time of each beat on the horizontal X axis, besides from the number of heartbeats represented on the graph line for each time. The time gap between each heartbeat is around 5 s.

7 Website for Management Control

To manage the application and present the gathered heartbeat data in a simple manner, a webpage was designed that enables the user to manage a new data entry (a), or view previously obtained data (b) while being in real-time connection, as shown in Fig. 4, and this is the last step of the process as it can be seen in Fig. 1e.

To begin a new heartbeat acquisition, it is necessary to provide the current user's name and begin the capture process by clicking on the start button, followed by the need to begin the capture by smartwatch application, as previously described. After that, the page displays information about the data for this new analysis, such as the

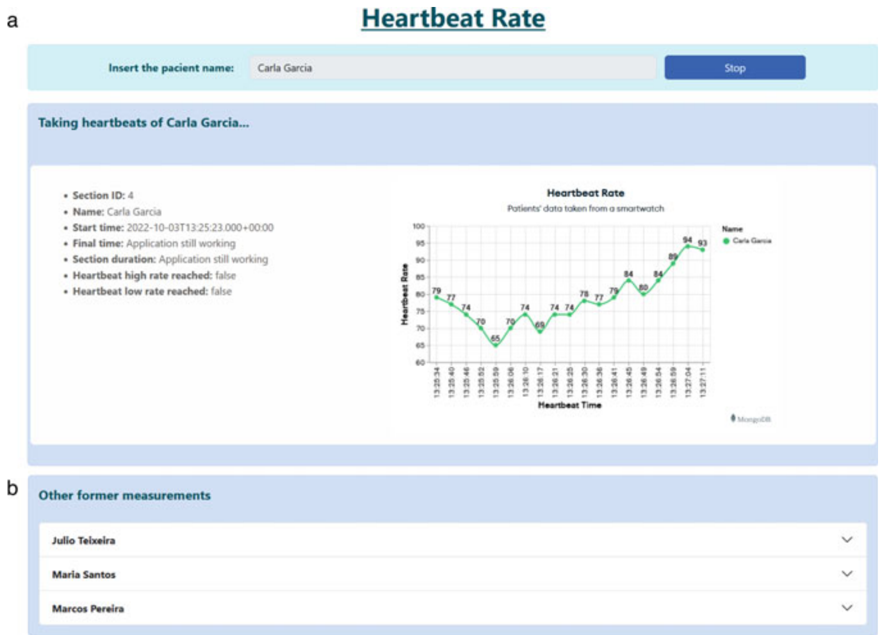


Fig. 4 Webpage exposure. **a** New entrance. **b** Old measurements

section identifier, name, start, finish, and duration of the capture section, and if the high or low limiting heartbeat rate was reached during application operation, showed on Fig. 4a. Along with this data, the chart automatically generates a real-time graph, displaying the heartbeat rate on the vertical Y axis and the time of each beat on the horizontal X axis, and the time gap between each heartbeat is approximately 5 s. Figure 4b displays the old measurements data, which is presented in the same manner as Fig. 4a.

8 Conclusions and Future Works

This research investigated the use of a smart device to capture heartbeats, saving data in a database and displaying it on a web page, all in real time and in a simple manner. Even though they are not medical devices, they perform very well in measuring vital signs and are able to aid in health sector in a way cheaper methodology compared to conventional medical equipment. It was realized that the cloud connectivity was quite useful, with the ability to record and visualize data in a few seconds, as well as transfer information by connecting devices to the database. With this implementation, the medical analysis of a schizophrenic can happen instantly while using the (VR) application.

The future of this implementation will enhance its database configuration, bringing in additional vital data such as health and national ID, as well as its presentation and features. Despite the good integration between the platforms used in the project's development, a limitation is the requirement to have the smartphone close to the watch and on the same account, which if not connected makes using the system impossible. Using outsourced services may also result in issues resulting from system updates or changes to privacy policies. This determines a potential change in the platforms used when focusing on the future growth of obtaining and displaying detailed vital signs of schizophrenics.

It will also be integrated with the other branches of the project that employ a (VR) serious game application in conjunction with a haptic vest to aid in the rehabilitation of schizophrenics and their reintegration into society. Heartbeat analysis will be crucial in determining how the patient feels during this treatment.

Also of interest are the additional possibilities for capturing vital signs that a smartwatch affords, which, if feasible, will be included to the project. Ideally, all the data obtained will be used to make our (VR) serious game application adapt to each single user, by the use of (AI), being possible to make the application take attitudes and change its approach based on the schizophrenic's body behavior.

Acknowledgements This work is funded by the European Regional Development Fund (ERDF) through the Regional Operational Program North 2020, within the scope of Project GreenHealth—Digital strategies in biological assets to improve well-being and promote green health, Norte-01-0145-FEDER-000042.

References

1. Abbas A, Yadav V, Smith E, Ramjas E, Rutter SB et al (2021) Computer vision-based assessment of motor functioning in schizophrenia: Use of smartphones for remote measurement of schizophrenia symptomatology. Digital Biomark 5:29–36. https://doi.org/10.1159/000512383
2. Alsaade FW, Alzahrani MS (2022) Classification and detection of autism spectrum disorder based on deep learning algorithms. Comput Intel Neurosci. https://doi.org/10.1155/2022/8709145
3. Lee PI, Hsueh PR (2020) Measurement of body temperature to prevent pandemic covid-19 in hospitals in taiwan–repeated measurement is necessary. J Microbiol Immunol Infect 53:365–367. https://doi.org/10.1016/j.jmii.2020.02.001
4. Rahaman H, Dyo V (2020) Counting calories without wearables: device-free human energy expenditure estimation. IEEE Comput Soc. https://doi.org/10.1109/WiMob50308.2020.9253424
5. Basjaruddin NC, Syahbarudin F, Sutjiredjeki E (2021) Measurement device for stress level and vital sign based on sensor fusion. Healthcare Inform Res 27:11–18. https://doi.org/10.4258/hir.2021.27.1.11
6. Abbas A, Yadav V, Smith E, Ramjas E, Rutter SB et al (2021) Computer vision-based assessment of motor functioning in schizophrenia: use of smartphones for remote measurement of schizophrenia symptomatology. Digital Biomark 5:29–36. https://doi.org/10.1159/000512383
7. Gupta S, Mahmoud A, Massoomi MR (2022) A clinician's guide to smartwatch 'interrogation'. Current Cardiol Rep 24(8):995–1009. https://doi.org/10.1007/s11886-022-01718-0
8. Nielsen AN, la Cour K (2022) Åse Brandt: feasibility of a randomized controlled trial of a proprioceptive and tactile vest intervention for children with challenges integrating and processing sensory information. BMC Pediatr 22. https://doi.org/10.1186/s12887-022-03380-5
9. Hirzle T, Gugenheimer J, Rixen J, Rukzio E (2018) Watchvr: exploring the usage of a smartwatch for interaction in mobile virtual reality. Assoc Comput Mach. https://doi.org/10.1145/3170427.3188629
10. Apple Inc. (2022) Apple watch series 6. https://www.apple.com/ge/apple-watch-series-6/
11. Apple Inc. (2022) Apple watch series 6 technical specifications. https://support.apple.com/kb/SP826
12. Apple Inc. (2022) Get the most accurate measurements. https://support.apple.com/en-us/HT207941
13. Apple Inc. (2022) Health data server. https://apps.apple.com/us/app/health-data-server/id1496042074
14. Apple Inc. (2022) Monitor your heart rate with apple watch. https://support.apple.com/en-us/HT204666
15. Novo A, Fonsêca J, Barroso B, Guimarães M, Louro A, Fernandes H, Lopes RP, Leitão P (2021) Virtual reality rehabilitation's impact on negative symptoms and psychosocial rehabilitation in schizophrenia spectrum disorder: a systematic review. Healthcare 9(11):1429. https://doi.org/10.3390/healthcare9111429
16. Lopes RP, Barroso B, Deusdado L, Novo A, Guimarães M, Teixeira JP, Leitão P (2021) Digital technologies for innovative mental health rehabilitation. Electronics 10(18):2260. https://doi.org/10.3390/electronics10182260
17. Github Inc. (2022) Health data server overlay. https://github.com/Rexios80/Health-Data-Server-Overlay
18. Github Inc. (2022) Virtual metro scenario for mental health rehabilitation. https://github.com/GreenHealthScholarship/Virtual-Metro-Scenario-for-Mental-Health-Rehabilitation
19. JSON: Json (2022) https://www.json.org/json-en.html
20. Lee PI, Hsueh PR (2020) Measurement of body temperature to prevent pandemic covid-19 in hospitals in taiwan—repeated measurement is necessary. J Microbiol Immunol Infect 53:365–367. https://doi.org/10.1016/j.jmii.2020.02.001

21. Lopes RP, Barroso B, Deusdado L, Novo A, Guimarães M, Teixeira JP, Leitão P (2021) Digital technologies for innovative mental health rehabilitation. Electronics 10(18):2260. https://doi.org/10.3390/electronics10182260
22. Meta Platforms I (2022) Oculus meta quest 2. https://www.meta.com/quest/products/quest-2/
23. MongoDB (2022) Mongodb. https://www.mongodb.com/
24. MongoDB (2022) Mongodb atlas. https://www.mongodb.com/atlas
25. MongoDB (2022) Mongodb chart. https://www.mongodb.com/docs/charts/
26. MongoDB (2022) What is mongodb. https://www.mongodb.com/en/what-is-mongodb
27. Nielsen AN, la Cour K (2022) Åse Brandt: feasibility of a randomized controlled trial of a proprioceptive and tactile vest intervention for children with challenges integrating and processing sensory information. BMC Pediatr. https://doi.org/10.1186/s12887-022-03380-5
28. Node.js (2022) Node.js. https://nodejs.org/en/
29. Novo A, Fonsêca J, Barroso B, Guimarães M, Louro A, Fernandes H, Lopes RP, Leitão P (2021) Virtual reality rehabilitation's impact on negative symptoms and psychosocial rehabilitation in schizophrenia spectrum disorder: a systematic review. Healthcare 9(11):1429. https://doi.org/10.3390/healthcare9111429
30. Quintero L, Papapetrou P, Munoz JE, Fors U (2019) Implementation of mobile-based real-time heart rate variability detection for personalized healthcare, pp 838–846. https://doi.org/10.1109/ICDMW.2019.00123
31. Rahaman H, Dyo V (2020) Counting calories without wearables: device-free human energy expenditure estimation. IEEE Comput Soc. https://doi.org/10.1109/WiMob50308.2020.9253424
32. Salekin A, Wang H, Williams K, Stankovic J (2017) Vrvisu—a tool for virtual reality based visualization of medical data. Institute of Electrical and Electronics Engineers Inc., pp 157–166. https://doi.org/10.1109/CHASE.2017.74
33. Winiarski S, Chomątowska B, Molek-Winiarska D, Sipko T, Dyvak M (2021) Added value of motion capture technology for occupational health and safety innovations. Human Technol 17:235–260. https://doi.org/10.14254/1795-6889.2021.17-3.4

DevOps Pragmatic Practices and Potential Perils in Scientific Software Development

Reed Milewicz, Jonathan Bisila, Miranda Mundt, Sylvain Bernard, Michael Robert Buche, Jason M. Gates, Samuel Andrew Grayson, Evan Harvey, Alexander Jaeger, Kirk Timothy Landin, Mitchell Negus, and Bethany L. Nicholson

Abstract The DevOps movement, which aims to accelerate the continuous delivery of high-quality software, has taken a leading role in reshaping the software industry. Likewise, there is growing interest in applying DevOps tools and practices in the domains of computational science and engineering (CSE) to meet the ever-growing demand for scalable simulation and analysis. Translating insights from industry to research computing, however, remains an ongoing challenge; DevOps for science and engineering demands adaptation and innovation in those tools and practices. There is a need to better understand the challenges faced by DevOps practitioners in CSE contexts in bridging this divide. To that end, we conducted a participatory action research study to collect and analyze the experiences of DevOps practitioners at a major US national laboratory through the use of storytelling techniques. We share lessons learned and present opportunities for future investigation into DevOps practice in the CSE domain.

Keywords DevOps · Scientific software development · Research software engineering

---

The first, second, and third authors contributed equally to this work.

R. Milewicz (✉) · J. Bisila · M. Mundt · S. Bernard · M. R. Buche · J. M. Gates · S. A. Grayson · E. Harvey · A. Jaeger · K. T. Landin · M. Negus · B. L. Nicholson
Sandia National Laboratories, Albuquerque, NM, USA
e-mail: rmilewi@sandia.gov

J. Bisila
e-mail: jbisila@sandia.gov

M. Mundt
e-mail: mmundt@sandia.gov

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_51

1 Introduction

High-performance computing (HPC) today plays a central role in scientific discovery, economic competitiveness, and national security in the United States and elsewhere. On the economic front, one US-government funded study estimated that every dollar invested in HPC generated an average of \$507 of new revenue and \$47 in profit or cost savings [1]. Likewise, to stay at the forefront of science and engineering, the United States has made substantial investments into computing technologies to push HPC into the Exascale era [2], with the world's first supercomputer capable of over a quintillion operations per second, *Frontier*, being brought online in 2022.

As the demand for high-quality simulations and data analyses continues to grow, the computational science and engineering (CSE) community has likewise had to evolve; there has been a notable shift away from small teams working on research scripts in isolation toward community-driven, open-source software ecosystems [3–6]. The makeup of the workforce has also been rapidly diversifying, with Research Software Engineering (RSE), DevOps, and IT Service Management (ITSM) professionals allying with computational scientists and mathematicians. They bring with them modern tools, practices, and perspectives on software development and maintenance – bridging the divide between conventional and scientific computing. Integrating those professionals into the teams, institutions, and culture remains an ongoing challenge [7–9], but the historical "chasm" between software engineering and scientific computing has narrowed considerably in recent years (*c.f.,* [10]).

In this study we focus our attention on DevOps in CSE contexts. For the purposes of this work, we use the definition of DevOps coined by Leite et al.:"a collaborative and multidisciplinary effort within an organization to automate continuous delivery of new software versions, while guaranteeing their correctness and reliability" [11]. To keep pace with demand, HPC CSE software must evolve more quickly while still remaining credible and trustworthy. There is an urgent need for more capable cyberinfrastructures to develop, deploy, and maintain that software. At the same time, however, DevOps for science and engineering presents unique challenges, and it demands adaptation and innovation in tools and practices; solutions that work for a web application in industry are unlikely to perfectly fit the needs of a multi-physics HPC application. At the present, the intersection of DevOps and scientific computing is critically understudied. A deeper understanding of how DevOps work is done in CSE contexts and what needs practitioners have could (1) inform the design of better tools and techniques, (2) support effective policy-making around cyberinfrastructure sustainment, and (3) raise awareness of the critical role played by DevOps practitioners in advancing science and engineering.

For these reasons, we conducted a participatory action research study to collect and analyze the experiences of DevOps practitioners at Sandia National Laboratories [12]. The first three authors of this study recruited practitioners to share "war stories" [13], detailed narratives of challenges faced and accomplishments made in DevOps work at the laboratories. In line with the principles of action research to allow software professionals to express their own voices, all participants were co-equally involved this study and are co-authors of this paper.

2 Background

DevOps has been in existence and usage for over a decade, evolving naturally from a necessity for breaking down silos between different developers within a software's lifecycle to focus on people and processes instead of distinct outcomes [14]. The shift is primarily marked by the creation of the *devopsday* conferences[1] in 2009 and has grown into a worldwide movement over the past 13 years. The emphasis of DevOps is to merge the "makers" with the "deployers" to create a more cohesive (and iterative) product; in fact, it is a natural extension of the Agile principles to extend beyond "code checkin" [15].

We can already see an issue with this emphasis, however—it is focused on creating and deploying a *product* in a pipeline where developers and operations professionals are not one and the same. Within scientific software communities, however, these activities have stayed essentially merged for research scientists. This is because, more often than not, researchers assume all of the roles within a software lifecycle [16]. Because software now underpins nearly all realms of scientific research, scientific researchers are expected to be literate not only in their domain of expertise, but also in software engineering.

There is a natural shift to respond to this need—including the application of DevOps practices to existing and new scientific software development teams. Through a combination of gray literature, peer-reviewed literature, and personal "war stories," we have observed that the adoption of these practices has large potential—but also potential pitfalls. We aim in this paper to discuss some of these successes and failures, contrasting how our stories compare with the top critical challenges to DevOps culture adoption as found in a recent systematic review by Khan et al. [17] and prioritized practices as discussed by Akbar et al. [18], to provide actionable suggestions for changes to the current paradigm, and to highlight areas where more research must be done.

3 Related Work

Almost all of the scholarly literature concerning DevOps for CSE has come from the perspective of researchers interested in applying DevOps tools and techniques rather than from DevOps practitioners. Conversely, there is a significant amount of gray literature (e.g., whitepapers, blog posts) that comes from the practitioners' perspectives.

In part, this is due to the recent up-trend of conferences and workshops geared toward scientific software development. Here is a non-exhaustive list of examples: the Collegeville workshop series[2]; the Tri-lab Advanced Simulation & Computing

[1] https://devopsdays.org/.

[2] https://collegeville.github.io/Workshops/.

Sustainable Scientific Software Conference[3]; and the Workshop on the Science of Scientific Software Development and Use[4]. Software sustainability has also been at the forefront of consideration for the Department of Energy as shown by the recent Request for Information on Stewardship of Software for Scientific and High-Performance Computing [19]. Cross-institutional projects such as the Exascale Computing Project (ECP)[5] also place importance on developer productivity and better development practices for scientific software, with information distributed through webinars, tutorials, and the website Better Scientific Software (BSSw)[6].

DevOps is a recurring topic in this space. In their BSSw blog post, Beattie and Gunter detail the adaptations of DevOps practices that have been applied to the Institute for Design of Advanced Energy Systems (IDAES), which include weekly standup meetings, incremental improvements to automated testing, and "soapboxing" (frequent discussions about the importance of software engineering practices with leadership) [20]. The 2020 Collegeville workshop's theme was "Developer Productivity" which yielded whitepapers that discussed Agile practices, challenges and successes related to automated testing, and a mapping of difficulties and recommendations for each stage of the software delivery lifecycle from the lens of scientific software engineering [21–24]. The 2022 Tri-lab Advanced Simulation & Computing Sustainable Scientific Software Conference had two tracks for "DevOps Infrastructure Development" and "DevOps CI/CD Pipeline Development."

de Bayser et al. has argued that DevOps concepts and practices should be integrated into the activities of researchers to help increase productivity and quality of the resulting software [25, 26]. Whitepapers and blogs from the DevOps community, however, argue for more specialized roles. Gesing argues for the implementation of well-defined roles in teams rather than researchers acting as "all-rounders" [27]. Adamson and Malviya Thakur second this view in their whitepaper on the operationalization of scientific software from a DevSecOps perspective [28]. This is further supported by the rise of the RSE professional designation which aims to represent the unique role of software engineering expertise applied directly into research software development.

4 Methodology

To collect and analyze the experiences of software practitioners doing DevOps work in CSE contexts, we used storytelling techniques to draw together an ensemble of challenges and triumphs in DevOps for CSE. We then analyzed that data through a participatory action research lens to build consensus among participants around their needs and values.

[3] https://s3c.sandia.gov/.

[4] https://web.cvent.com/event/1b7d7c3a-e9b4-409d-ae2b-284779cfe72f/summary.

[5] https://www.exascaleproject.org/.

[6] https://bssw.io.

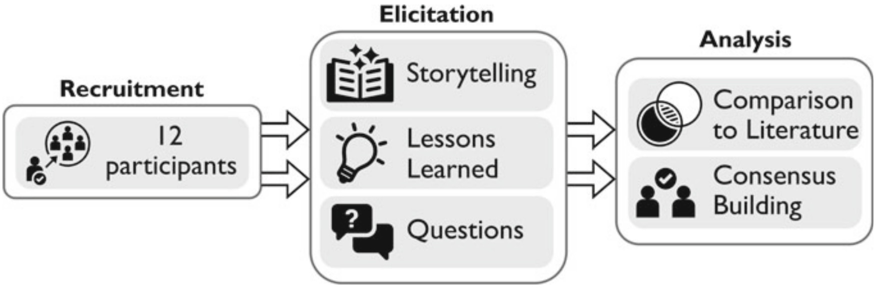


Fig. 1 An illustration of the methodology used to collect and assess evidence gathered in our study. Participants were recruited to share stories of challenges and triumphs in DevOps, report lessons learned, and reflect on their questions about best practices. For our analysis, we compared reported experiences to trends in the scholarly literature and then iterated with our participants as co-researchers to refine the contents of this study.

Storytelling is a qualitative data collection technique where participants are asked to recount detailed events from their own experiences [29]. The use of storytelling to analyze participants' experiences is a popular technique in the social sciences. A 2022 publication of *Communications of the ACM*, Barik et al. speak to the importance of applying storytelling techniques in scientific settings to promote better communication and overall understanding [30]. Moreover, as noted by Polletta et al., as a method for representing the views, attitudes, and experiences of a community, storytelling is often seen as more authentic and democratic in character (*i.e.*, "everybody has a story") [31].

Our study draws inspiration from the work of Lutter and Seaman, who collected "war stories" concerning documentation usage during software maintenance [13]. A key methodological difference in our work is that we seek to apply methodological techniques from participatory action research (PAR) to guide our data collection and analysis. PAR is an approach to action research that emphasizes direct participation in the research process by the members of the community whose interests the research is meant to serve [32]; as explained by Baum et al., "PAR advocates that those being researched should be involved in the process actively" [33]. In our work, we take the position that software practitioners doing DevOps for CSE are qualified subject matter experts who can speak credibly to the challenges they face, and that all participants in our study (the first three authors included) are co-researchers. For that reason, rather than collecting data from participants and independently performing qualitative analysis on that data, we used an iterative, consensus-based approach to draw out themes among our experiences. To help lend greater validity to the work and mitigate bias, we draw upon the peer-reviewed literature to compare and contrast our experiences with those of other DevOps practitioners (see Fig. 1).

All the authors of this study are employed at Sandia National Laboratories, a US federally-funded research and development center (FFRDC)[7]. As national secu-

[7] https://www.sandia.gov.

rity laboratory, Sandia relies heavily on computational simulation and data analysis to achieve its science and engineering objectives; this is made possible through a complex ecosystem of scientific software libraries and applications, some developed internally and others community-owned and hosted on the open web. Orchestrating the development, deployment, and maintenance of those software stacks is a significant DevOps research and development (R&D) challenge and an active area of interest for US national laboratories. During the Fall of 2022, the first three authors recruited participants through the institution's Research Software Engineering Community of Practice (RSE-COP) mailing list and directly from the first three authors' departments (roughly 150 people in total). Participants confirmed their participation over email and submissions for stories were collected using a collaboratively-edited, web-based corporate wiki. The contribution page included guidelines for potential contributors.

In particular, potential contributors were asked to provide stories about their experiences conducting DevOps in scientific software development (successes, failures, challenges, changes, etc.) in topics such as testing, team policies and procedures, technology stack modernizations, tradeoffs (e.g., maintainability for performance), etc. To seed the discussion, the first three authors provided stories and open questions. Recruitment yielded 9 additional participants, and together we produced a set of 13 stories that reflect different aspects of DevOps work at the labs. In addition to the stories, contributors added open questions relating to their story or the DevOps culture and ecosystem. We present those stories and questions in Sect. 5.

To better understand how our experiences map to those of DevOps practitioners outside of CSE contexts, we analyze their challenges and lessons learned through the lens of the scholarly literature on DevOps in industry contexts. In particular, we use two systematic reviews of the literature to frame our analysis: a review of common cultural challenges to DevOps adoption in organizations by Khan et al. [17] and a review of best practices in DevOps by Akbar et al. [18]. We present findings from our analysis in Sect. 6.

5 Results

Following the collection of the stories, we identified four overarching themes: Software Development Lifecycle, Testing, Team Policies and Processes, and Institutional Support. We present the results of those themes here, including open questions posed by the authors, and will discuss them in more detail in Sect. 6.

5.1 Software Development Lifecycle

Software development of all forms will execute, whether explicitly or implicitly, a software development lifecycle (SDLC) model [34]. Popular models are Agile, Kanban, and (the topic of this paper) DevOps. At their core, SDLC models provide defined structure for software development activities.

CSE teams also employ SDLC models, such as Use Case Driven Development. “Use Cases” are the fundamental piece of business value in most software. They can help clearly communicate the business needs from the customer to the development and research staff. Use Case Driven Development is a methodology focused on using “Use Cases” as the central component to writing software [35].

The application of this methodology to scientific software development can be quite natural. As one author details:

> At the start of a new project, we knew that the customer would have a set of research questions related to their data and domain. To design the software, we modeled each research question as a use case and had a few planning meetings to further define entities and relationships. ... Each use case became a command line tool and separate python module that the customer could use directly. We developed a common set of libraries that define and work with the domain entities and relationships—**S1**.

In practice, this author found that applying Use Case Driven Development strategies to their research resulted in much more cohesive conversations around and development of the research software. Each feature could be directly mapped to a “Use Case,” and because of the modularity of the design, it became much easier to augment when new questions were added to the domain.

While Use Case Driven Development is useful for creating software, there is still the consideration of deploying it. A common challenge particularly across national laboratories is how to ensure cohesive usage across differing customers, networks, computing systems, and architectures. Two of the authors describe their solutions to this challenge:

> We have started using Docker containers to rapid and flexibly deploy software to our customers. ... Using this set of Docker containers removes the concern about customizing an environment on the customer machine(s). Instead, we can customize everything on our end and then send them the set of Docker containers as a zip file. In practice this has increased our success rate of deployment and allowed us to ensure that our development environment is nearly identical to our deployment environment—**S2**.

While this solution is tuned to an external customers’ needs, another of the authors instead looked at how to resolve differing build environments within the same development team:

> One of the consistent workflow problems that my colleagues and I have run into, especially when on-boarding new team members, is to get someone set up with the proper development environment, the correct versions of libraries, etc. for a particular project. ... In the last few years, a number of my colleagues and I have managed our build environments using the Nix system, and it has been quite beneficial to our projects. When on-boarding new people, all they have to do is execute the command `nix develop` in the root of the source tree. Occasionally there is a hiccup, but the vast majority of the time they get a fully configured development environment without any additional effort. This saves us days or weeks of frustration—**S3**.

These stories detail successful application of industry-standard DevOps solutions to scientific software development needs.

5.2 Testing

The challenges surrounding testing scientific software have been well-documented. In Kanewala and Bieman's literature review, they detailed that these challenges come in two forms: technical and cultural [36]. The authors have experienced the problems in both of these major categories.

Technical In the category of technical challenges, Kanewala and Bieman further categorize into four sections: (1) test case development, (2) producing expected test case output values, (3) test execution, and (4) test result interpretation. Independent of this literature review, the authors provided stories (some failures, some successes) that fall into each of the four sub-categories.

With regards to (1), one author specifically calls out difficulties relating to number of potential parameters:

> Some bugs were able to slip past our CI/CD. These bugs were usually missed because the CI/CD did not fully exercise the parameter space (e.g., build options). The team is looking to fix this by either running a full factorial parameter matrix or by decomposing the behavior into independent units which can be tested independently rather than compositionally—**S4**.

Another author also alludes to configuration options in their testing infrastructure:

> I am member of a software package that consists of dozens of sub-packages with thousands of configuration options, all of which application teams rely on for their specific sub-package. Continuous integration testing entails vetting that the code-base works with specific configuration settings; toolchains such as GCC, CLANG, and CUDA; and HPC architectures—**S5**.

The same author also details challenges with regards to (3):

> The CI testing infrastructure is currently limited to using a custom automation tool that pulls the proposed code changes into [our institution's] networks. The tool must then launch and monitor the tests. With build and test times averaging six hours and up to 11 builds run per change, there is a huge maintenance and resource cost. Frequently, something goes wrong during the average build and results in test results never being reported back to the developer—**S5**.

While not explicitly stated within (3), modernization of testing infrastructures was another consideration of the authors with respect to test execution.

> I am a DevOps contributor to a scientific Python package aimed at optimization modeling. ... In early 2022, a downstream dependency requested support for Python 3.10. This request revealed a problem ... [that] required an entire refactor of the testing suite to use the popular and regularly maintained package pytest. On the surface, this refactor seemed simple. The issue: because of the age of the scientific package, ample homegrown infrastructure had been built specifically around nosetests that needed to be preserved (e.g., dynamic categorization, dynamic test creation). What was anticipated to be a quick and simple fix took over a month of dedicated, time-intensive work to convert in order to maintain expected functionality—**S6**.

Multiple authors have had to contend with one of the largest plagues in scientific software: repeated code (4).

> Often times in engineering science software libraries, there are many similar implementations (models, etc.) that could share a lot of the same core tests. This includes integration tests in addition to unit tests. The problem is that the many similar implementations rarely share the same tests, since they are typically developed in series rather than in parallel. ... This causes issues related to inconsistent testing of implementations when certain tests are included somewhere and not elsewhere, and a lack of testing efficiency when the tests are copied in multiple places. Theses issues can inhibit the quality of the software and cause further development to become less straightforward—**S7**.

In some cases, the repeated code led to the developers struggling with the same bug for over a decade:

> While running an important application deployed to an HPC cluster, it was discovered that there was a scalability bug in a math library we develop, resulting in a nearly 30% drop in performance. ... What was interesting, however, was the unusually long-lived history of the bug. The team of developers who found the bug discovered that the exact same bug had been introduced, found, and fixed in the math library multiple times over the years. The offending code was first introduced in three packages between 1998–2000 and fixed in 2005, copied line for-line into a fourth package in 2004 and fixed again in 2015, and finally introduced into the last package in 2014 and fixed in 2017. In each case, the discovery and solutions were socialized, comments were made in the code, and notes were left in an issue tracker, but that information did not flow to the right parties in each subsequent incident—**S8**.

While the DevOps challenges surrounding testing in scientific software development are well-known and commonly experienced, testing has long-reaching implications on the success and stability of the software. One author describes the benefit of formal verification (2):

> I regularly contribute to a large code base in a domain that is notorious for begetting intricate software systems that contain all sorts of subtle bugs, some which can live for decades. The particular code that I contribute to is unique because it has a formal, end-to-end, proof of correctness, which is mechanically checked against the code at compile-time. [In o]ne of the improvements that I worked on ... I hit a wall in the proof and could not proceed further. When investigating why the proof wouldn't work, I realized that my implementation was incorrect. I had missed a corner case. In the end, I might have been able to catch this corner case with tests, but the application domain is complex enough where that bug could easily have gone unnoticed. I was able to fix this bug before it ever made it into the code-base—**S9**.

Some of these experiences, however, crossed the line from technical into cultural.

Cultural Examples of technical challenges experienced by the authors are endless. They are not, however, strangers to the cultural concerns as well. While story **S9** talks about the benefit of formal verification, the author also notes a significant problem—formal education:

> The overwhelming majority of scientists and software engineers have absolutely no experience in formal verification. It just seems like a black art. How do we educate our workforce about practical formal verification? Just as with "design-for-test," how do we integrate "design-for-verification" into our programming curricula?—**S9**

One author calls out specific ways that education can be applied:

> The greatest difficulty I've had as a DevOps practitioner in the research software world has been getting decision makers to understand the complexity involved in designing, building,

> maintaining, and extending infrastructures to accelerate the delivery of value from development into operations, amplify feedback loops, and enable a culture of continual learning and experimentation. Building shared understanding is a prerequisite for culture change. Applying this to DevOps in the research software space means algorithmists, simulations experts, analysts, managers, etc., must dedicate time away from regular milestone-driven activities to learn what the DevOps paradigm shift actually is, what changes in thinking it requires, and what kind of activities it entails. This can be done, in part, through studies and discussions of books such as *The DevOps Handbook*, *The Phoenix Project*, *The Unicorn Project*, or *Continuous Delivery*—**S10**.

This author points out a secondary problem—software development activities fall below research priorities. In reference to story **S6** regarding the conversion from `nosetests` to `pytest`:

> As software projects mature, so, too, should their support for modern technology. In this case, once the main test driver was announced to no longer be supported or updated, it would have benefited the scientific software development team to begin the transition to a newer, regularly maintained test driver. Instead, the team relied on the hope that it would continue to work... Until it didn't. It is essential for teams to preemptively address these concerns rather than wait—**S6**.

The overarching consensus for cultural issues throughout these stories: the expertise of the DevOps developers needs to be given the same priority as those of the domain scientists.

5.3 Team Policies and Processes

Software quality is dependent on the effectiveness of a project's DevOps practices, but this goes further than just quality—culture also plays a critical role [37]. As stated by Perera et al.: "Culture is another important factor because it changes the way in which teams work together and share the responsibility for the end users of their application."

As detailed above, one author contributes to a package with dozens of subpackages and thousands of configuration combinations. In their case, the difficulties in testing lead directly to poor teaming dynamics:

> Developers are frustrated while the DevOps team has an endlessly growing backlog of work. While this automation is better than no CI testing, it has resulted in poor teaming dynamics and a large maintenance burden. Additionally, this DevOps team is so busy maintaining configurations and keeping the infrastructure running that they have no time (or available options) to improve the infrastructure. As a result, teaming dynamics continue to digress, and it is difficult to retain team members—**S5**.

Teaming dynamics can be strained more with changing policies and procedures, though the risk may pay off in the long run. One author details their package's shift from subversion to git and hosting on GitHub:

> The more mature a project is, the more likely there has been turmoil over changing technologies. For example, one package created by [laboratory] scientists started on subversion

> and a [internally]-hosted repository for many years before transitioning to GitHub. In the subversion days, developers would commit directly to the main branch, which led to frequent bugs, breakage, or repository pollution. ... When the team eventually transitioned to GitHub, pull requests and code reviews were added into the development workflow. Initially this caused conflict on the team as it slowed down the development speed and introduced "extra overhead." This change, however, improved the overall stability of the code base. ... The code reviews also generally have raised the quality of the code base. This has allowed developers and maintainers to shift focus to improving existing infrastructure and modernizing the code and its dependencies and has allowed a larger community of contributors to add their contributions without lowering the quality of the package—**S11**.

This example highlights two main points: (1) change can be difficult but overall bring about better processes and stability, and (2) buy-in is essential for adopting team policies. Once fully adopted, solid team policies can save a team from disaster. Another author provides a perfect example:

> A graduate student intern collaborating with [laboratory] researchers was traveling to a customer meeting to demonstrate their software product, including very recent updates to the tool and presentation. During travel, the student's laptop—the primary machine used for development—was stolen. Fortunately, members of the team had diligently maintained comprehensive remote version control systems. Upon arrival at the meeting location, the presentation, software tool, and demonstration were downloaded to a colleague's system, reverted to a stable version, and operations proceeded with minimal disruption. Customers who were unfamiliar with the team's version control practices were thoroughly impressed at the team's resilience given the circumstances—**S12**.

This same author goes further to say, "Incorporating DevOps best practices into workflows hedges against unforeseen catastrophe. The initial investment and learning curve associated with applying these strategies routinely—especially for scientists who may feel little need to otherwise learn "software development" skills—has the potential for serious payoff in the long run."

5.4 Institutional Support

Providing institutional support for development opportunities (professional and technical) is a key contributor to retention of staff and staff happiness [7]. In particular, providing opportunities for staff recognition, growth, and ability to influence the organization's direction through implementing cultural changes, teaching new best practices, and advancing an organization's shared understanding of DevOps best practices can boost retention and create a sense of belonging.

Professional development opportunities can come in different forms: taking training, developing training, community building, etc. There has long been a history of software developers within CSE teams being fragmented in their work [38]. This type of fragmentation primarily affects access to professional development opportunities that would improve the overall state and trustworthiness of scientific software. As one author points out:

> The problem is we have a long history of insufficiently funding and staffing the [software development] activities that would solve our issues and prevent them from happening again in the future—**S10**.

In their study, Raybourn et al. found much of the same: "The next opportunity area for incentivizing software quality engineering as part of a culture's practice comes in the allocation of funding for quality in software projects. In interviews with personnel from two distinct Centers, lack of monetary resources was the most-frequently stated challenge facing developers. Participants mentioned fragmentation, competition for funding, lack of rewards for development work, etc. As one participant plainly stated, "You can't have quality without the money to pay for it."" [39]

Institutional support is necessary not only for direct development activities, but also training. One such author was fortunate to be financially supported in creating and delivering a training for Git:

> Running a 2-hour long workshop and teaching others about Git was a first for me. Once the reality sank in and I was getting ready for the talk and the hands-on activity, I realized that I didn't know the tool as well as I ought to to be able to present to others with any sense of authority. I asked around for resources and compiled the important topics to cover. From practicing Git for about a decade now, I knew how to use it and navigate its documentation to get the work done, but I did not feel comfortable answering questions on the spot about its nuances. Being confronted with my lack of knowledge forced me to better my understanding of Git and it allowed me to become a more proficient user—**S13**.

This activity provided the author with the ability to not only inform others, but also strengthen themselves as a developer.

5.5 Open Questions

- How can we document and standardize our deployment process to make it even easier to release and deploy software for researchers without DevOps training?
- How do you find the time/funding to make such drastically large changes?
- Is it possible to get funding agencies to start requiring technical debt reduction plans as part of the proposal process? How do we promote a culture of mutual ownership of technical debt?
- How do we build a culture among those who identify as scientists first that software best practices should be adopted and adhered to, even when the initial investment is high?
- How do you appropriately allocate time and resources for making small incremental changes in an effort to avoid technical DevOps debt accumulation and poor teaming dynamics?
- Can knowledge sharing among development teams be encouraged by creating formal roles for people doing DevOps work? What would that look like?
- Are there tools that can help with code clone detection at production-scale?
- Is there a well-defined process that already exists (code agnostic, of course) that tackles the problem of duplicate tests?

6 Analysis

The stories recorded above offer a window into DevOps practice at a major scientific institution. In this section, we seek to ground those experiences in the scholarly literature around DevOps in industry to draw out similarities and differences to DevOps in CSE contexts. In Table 1, we provide summaries of the key lessons learned in each story; these recommendations may be valuable for computational scientists, engineers, and those doing DevOps work in CSE contexts.

We first compare the challenges faced and addressed by our participants to those commonly attested in the scholarly literature on DevOps practice. Using Khan et al.'s systematic review of challenges to adoption of a DevOps culture as a guide, we found support in our stories for eight out of the ten most-frequently mentioned challenges in the literature (see Table 2). That is, many of the obstacles encountered and overcome by our participants are not unique to CSE software development, which lends credibility to the generalizability of our findings and recommendations. Successful implementation of DevOps tools and practices requires addressing quality in the systems-under-test and the tests themselves (S4, S7, S9, S13), and a lack of knowledge about DevOps can hinder implementation (S9, S10, S12) or lead to suboptimal decision-making (S1, S2). Where DevOps infrastructure exists, it must be actively maintained to keep it up-to-date (S6, S11) and to manage complexity creep (S3, S5). Having the support of leadership and buy-in from development teams (S10, S11) is critical to success, as is promoting communication and collaboration across teams and organizations (S8).

Next we compare the successes and better practices experienced by our participants to Akbar et al.'s "prioritization-based framework of the DevOps best practices based on evidence collected from industry experts." We found support in our stories for ten of the twelve top-ranked practices (see Table 3). It is worth noting that Akbar et al. divide their rankings into global vs. local ranks (e.g., overall highest priority vs. priority within a common category). The top ten globally *all* fall within the "Culture" category; we opt to use the top three ranked practices in each local category instead to diversify the conversation. As shown, besides two relatively industry-centric prioritizations ("microservices" and "tools to capture requests"), all of the other top-ranked practices are reflected in the stories. DevOps adoption and implementation is seen as successful within a more collaborative culture with a shared value system (S8, S10, S11, S13), which is bolstered by the education of both staff and leadership (S12). The practices succeed only insofar as there is standardized buy-in and adherence (S1, S8, S10, S11), and constant communication is seen as necessary to minimize the potential for issues and inefficiencies (S5, S8, S10, S13). Only one story touches on rapid deployment as a way to receive constant feedback from customers (S2). When it comes to automation, there is a strong emphasis on the importance of continuous integration and testing (S4, S5, S6, S7, S9), but this cannot be applied unless a team can decide what it wants to actually achieve first (S1, S6, S11). Taking this one step further, DevOps practices are seen as overall more effective when adopted early while a project is still small (S8, S10, S11, S12) and are more successful with

Table 1 Summaries of lessons learned from the stories generated during the storytelling exercise.

Story	Topic	Summary
S1	*Align Tools and Methodologies*	Teams can be highly productive when their DevOps tooling matches with and supports their development methodology.
S2	*Embrace Virtualization and Interactive Media for Deployment*	Technologies like containerization (e.g., Docker) and interactive computing media (e.g., Jupyter notebooks) enable projects to rapidly deploy CSE software to customers.
S3	*Adopt Dependency Managers*	As CSE software ecosystems continue to grow and mature, emerging dependency management solutions like Nix and Spack can help manage that complexity.
S4	*Use a Mix of Testing Strategies to Ensure Code Quality*	While CI/CD is effective at catching certain kinds of bugs, it is not a panacea. It is important to test across different environments and configurations with different tiers of testing and to design software to be more testable.
S5	*Maintain Good Tooling to Have Good Teaming*	Inefficient, fragile, and complex infrastructure drains the energy and morale of DevOps practitioners. Refreshing and incrementally improving infrastructure is vitally important for effective teaming.
S6	*Manage Test Infrastructure*	Teams must proactively maintain their build and test infrastructure, as it is guaranteed to fail eventually.
S7	*Design Software for Testing*	CSE software modules and components should be built with testing in mind, such as by having common interfaces against which tests can be written.
S8	*Address Human Communication Bugs*	As projects scale up, it becomes essential not only to put in place DevOps infrastructure but also to build out processes and practices to facilitate coordination between teams.
S9	*Leverage Static Analysis and Formal Methods*	Some of the worst, most complicated software bugs can be prevented through automated software analyses, ranging from built-in type-checking to robust theorem proving tools.
S10	*Promote Shared Understanding for Culture Change*	For DevOps efforts to succeed, there needs to be a shared understanding throughout the organization of what the DevOps paradigm shift actually is, what changes in thinking it requires, and what kind of activities it entails.
S11	*Build Consensus Around DevOps Tools and Practices*	Mature projects accrue inertia around existing tools and practices, and securing buy-in on new tools and practices is essential for adoption to succeed.
S12	*Teach DevOps Practices and Principles*	DevOps practitioners should educate their peers on best practices, both to reinforce their own knowledge and to improve the state of practice in their institutions.
S13	*Plan for Resiliency*	Getting CSE teams to adopt DevOps tools and practices helps guard against unexpected catastrophes, and this should be emphasized as a benefit of having those tools and practices in place.

Table 2 Common cultural challenges to implementing DevOps in organizations (as identified by Khan et al. [17]) attested and/or mitigated in the stories we collected.

Common cultural challenges	Discussed/mitigated in stories
Lack of collaboration and communication	S8
Lack of skill or knowledge about DevOps	S9, S10, S12
Culture of blame (criticism)	–
Lack of intentional DevOps approach	S1, S2
Lack of management support	S10
Trust and confidence issues	S10, S11
Complicated infrastructure	S3, S5
Poor quality	S4, S7, S9, S13
Security issues	–
Legacy infrastructure	S6, S11

Table 3 Top-ranked DevOps practices (as identified by Akbar et al. [18]) mirrored in the stories we collected

Prioritized cultural practices	Discussed in stories
Collaborative culture with shared goals	S8, S10, S11, S13
Readiness to utilize a microservices architecture	–
Education of executives	S12
Prioritized sharing practices	Discussed in stories
Standardized processes and procedures	S1, S8, S10, S11
Continuous feedback to address issues and inefficiencies	S5, S8, S10, S13
Reduce batch size to increase communication	S2
Prioritized automation practices	Discussed in stories
Decide what to do first	S1, S6, S11
Continuous integration and testing	S4, S5, S6, S7, S9
Use tools to capture every request	–
Prioritized measurement practices	Discussed in stories
Effective and comprehensive measurement/monitoring	S4, S9
Start DevOps on small projects	S8, S10, S11, S12
Integrated configuration management	S2, S3

frequent monitoring and adaption of practices as a project matures (S4, S9). They can also lead to better ease-of-use when there is a cohesive and integrated system for managing dependencies and environments (S2, S3).

Overall, we see a trend that DevOps practices and perils align across industry and CSE contexts—however, within CSE, there are necessary adjustments. For example, while our participants agreed with the necessity for standardized processes and procedures, the scope differs. In industry, this standardization is desired across the entire organization; for CSE, it's enough for the standardization to be across a project or projects that collaborate. They also do not need to be completely standardized; rather, they should be appropriately scaled as funding and expertise allow. CSE teams can succeed by applying appropriately scaled practices to their projects, but they must avoid the perils that come with either trying to do too much or too little.

Open questions remain surrounding how to do this adaptation in a scaleable, reproducible manner. For CSE projects that start as completely exploratory and state-of-the-art, at what point does the team "scale up" their DevOps practices? What tools and methods for continuous integration should be applied at each stage of maturity? How often should a research code release and deploy, and what does operations and maintenance really look like after the fact? These are all potential future avenues for research.

7 Threats to Validity

As with every study, this one has its threats to its validity. We will discuss three such potential threats here: (1) generalization; (2) qualitative nature of the data; and (3) personal biases.

The authors (and participants) in this study all come from the same institution and have similar job types. This presents a threat in being able to widely generalize the results. In particular, we cannot say for certain that all DevOps practitioners who collaborate with CSE teams will have these same views; however, we have aimed to mitigate this by pairing the authors' experiences with support from peer-reviewed and gray literature.

As for the second threat, all data presented here is purely qualitative. It can be difficult to establish concrete trends and conclusions; however, as with all studies, we consider this a trade-off. For this study, we believe the stories encapsulate a fuller, richer, and more complete picture of what DevOps work looks like in practice. Additionally, similar to the previous threat, we have attempted to mitigate this with peer-reviewed and gray literature to provide more breadth of experience to our own.

As a final note, we want to call out the potential of our own personal biases. Because the authors themselves are the participants, we recognize the potential to skew the results based on our assumptions and feelings rather than actual fact. To mitigate this, we collectively reviewed perspectives and content contributed by each other. In this way, we aimed to minimize possible bias while still preserving the expressed opinions.

8 Conclusion

The DevOps movement may have its roots in industry, but it has branched into scientific software development. As software becomes more integral to the advancement of science, so do the processes and procedures used to create scientific software.

In this article, the twelve authors shared thirteen unique stories of their experiences as DevOps practitioners within CSE teams within Sandia National Laboratories, including lessons learned and residual open questions. Using a participatory action research approach, we combined these stories with gray and peer-reviewed literature to analyze the commonalities and differences between industry and scientific software DevOps practices.

We found that many practices and perils are mirrored. CSE teams experience the same cultural challenges as industry while emphasizing similar priorities on testing, collaboration, and starting early. With that in mind, DevOps practices cannot be perfectly applied out-of-the-box to a CSE project. There needs to be adaptation, education, and buy-in to create success.

DevOps in the CSE context is a research realm that is rich in unanswered questions. We shared some of these, as well as lessons learned, to add to and promote further conversation around pragmatic practices and potential perils in scientific software development.

Acknowledgements Illustrations used with permission from thenounproject.com by Kamin Ginkaew, Adrien Coquet, Numero Uno, and Gregor Cresnar.
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. SAND2022-16099 C.

References

1. Joseph E (2020) SC20 update on the ROI and ROR from investing in HPC. https://www.hpcuserforum.com/ROI/
2. Douglas K, Stephen L, Irene Q (2018) Exascale computing in the United States. Comput Sci Eng 21(1):17–29
3. Wilhelm H, Leslie C, Simon H, Heather P, Thanassis T (2020) Open source research software. Computer 53(8):84–88
4. Heise C, Pearce JM (2020) From open access to open science: the path from scientific reality to open scientific communication. SAGE open 10(2):2158244020915900
5. Arne J, Wilhelm H (2018) Software engineering for computational science: past, present, future. Comput Sci Eng 20(2):90–109
6. Tennant JP, Agrawal R, Baždarić K, Brassard D, Crick T, Dunleavy DJ, Evans T R, Gardner N, Gonzalez-Marquez M, Graziotin D et al (2020) A tale of two 'opens': intersections between free and open source software and open scholarship
7. Mundt M, Beattie K, Bisila J, Ferebaugh C, Godoy W, Gupta R, Guyer J, Kiran M, Malviya-Thakur A, Milewicz R, Sims B, Sochat V (2022) For the public good: connecting, retaining, and recognizing current and future RSEs at national organizations (under review). In: Special issue on the future of research software engineers in the US. Comput Sci Eng

8. Mundt M, Milewicz R (2021) Working in harmony: towards integrating RSEs into multidisciplinary CSE teams. In: Proceedings of the 2021 workshop on the science of scientific-software development and use. U.S. Department of Energy, Office of Advanced Scientific Computing Research
9. Mundt M, Sochat V, Katz DS, Gesing S, Melessa Vergara VG (2021) DOE software stewardship challenges in diversity, professional development, and retention of research software engineers. In: Responses to the request for information on stewardship of software for scientific and high-performance computing. U.S. Department of Energy
10. Kelly DF (2007) A software chasm: software engineering and scientific computing. IEEE Softw 24(6):120–119
11. Leonardo L, Carla R, Fabio K, Dejan M, Paulo M (2019) A survey of DevOps concepts and challenges. ACM Comput Surv (CSUR) 52(6):1–35
12. Chevalier JM, Buckles DJ (2019) Participatory action research: theory and methods for engaged inquiry. Routledge, Milton Park
13. Lutters WG, Seaman CB (2007) Revealing actual documentation usage in software maintenance through war stories. Inform Softw Technol 49(6):576–587
14. Davis J, Daniels R (2016) Effective DevOps: building a culture of collaboration, affinity, and tooling at scale. O'Reilly Media, Inc
15. Mueller E (2010) What is DevOps?
16. Bernholdt DE, Cary J, Heroux M, McInnes LC (2021) Position papers for the ASCR workshop on the science of scientific-software development and use. Technical report, USDOE Office of Science (SC)
17. Khan MS, Khan AW, Khan F, Khan MA, Whangbo TK (2022) Critical challenges to adopt DevOps culture in software organizations: a systematic review. IEEE Access 10:14339–14349
18. Azeem Akbar M, Rafi S, Alsanad AA, Furqan Qadri S, Alsanad A, Alothaim A (2022) Toward successful DevOps: a decision-making framework. IEEE Access 10:51343–51362
19. Hal F, Ben B, Robinson P, Saswata H-M, Bill S (2021) Responses to the request for information on stewardship of software for scientific and high-performance computing. Technical report, USDOE Office of Science (SC)
20. Beattie K, Gunter D (2021) Useful practices for software engineering on medium-sized distributed scientific projects
21. Dubey A (2020) When not to use agile in scientific software development. In: The 2020 Collegeville workshop on scientific software
22. Ellingwood N, Rajamanickam S (2020) Practices and challenges of software development for a performance portable ecosystem. In: The 2020 collegeville workshop on scientific software
23. Finkel H (2020) The many faces of the productivity challenge in scientific software. In: The 2020 Collegeville workshop on scientific software
24. Windus T, Nash J, Richard R (2020) Scientific software developer productivity challenges from the molecular sciences. In: The 2020 Collegeville workshop on scientific software
25. De Bayser M, Azevedo LG, Cerqueira R (2015) ResearchOps: the case for DevOps in scientific applications. In: 2015 IFIP/IEEE international symposium on integrated network management (IM), pp 1398–1404
26. de Bayser M, Segura V, Azevedo LG, Tizzei LP, Thiago R, Soares E, Cerqueira R (2022) DevOps and microservices in scientific system development: experience on a multi-year industry research project. In: Proceedings of the 37th ACM/SIGAPP symposium on applied computing, pp 1452–1455
27. Gesing S (2020) Increasing developer productivity by assigning well-defined roles in teams. In: The 2020 collegeville workshop on scientific software
28. Adamson R, Malviya Thakur A (2021) Perspectives on operationalizing scientific software. In: The 2021 collegeville workshop on scientific software
29. Patton MQ (2002) Qualitative research & evaluation methods. Sage, Newcastle upon Tyne
30. Titus B, Sumit G, Mario J (2022) Storytelling and science. Commun ACM 65(10):27–30
31. Polletta F, Chen PCB, Gardner BG, Motes A (2011) The sociology of storytelling. Ann Rev Sociol 37(1):109–130

32. Reason P, Bradbury H (2008) Handbook of action research: participative inquiry and practice. Sage, Newcastle upon Tyne
33. Fran B, Colin MD, Danielle S (2006) Participatory action research. J Epidemiol Commun Health 60(10):854
34. Ralf K (2017) Sixty years of software development life cycle models. IEEE Annal Hist Comput 39(3):41–54
35. Alexander IF, Beus-Dukic L (2009) Discovering requirements: how to specify products and services. Wiley, New York
36. Kanewala U, Bieman JM (2014) Testing scientific software: a systematic literature review. Inform Softw Technol 56(10):1219–1232
37. Perera P, Silva R, Perera I (2017) Improve software quality through practicing DevOps. In: 2017 seventeenth international conference on advances in ICT for emerging regions (ICTer), pp 1–6
38. Katz DS, McHenry K, Reinking C, Haines R (2019) Research software development & management in universities: case studies from Manchester's RSDS group, Illinois' NCSA, and Notre Dame's CRC. In: 2019 IEEE/ACM 14th International Workshop on Software Engineering for Science (SE4Science), pages 17–24. IEEE, 2019
39. Raybourn E, Milewicz R, Mundt M (2022) Incentivizing adoption of software quality practices. Technical report, Sandia National Laboratories (SNL-NM), Albuquerque, NM

Mining Fleet Management System in Real-Time "State of Art"

Hajar Bnouachir, Meriyem Chergui, Mourad Zegrari, and Hicham Medromi

Abstract The objective of this paper is to understand real-time fleet management problems, its characteristics, and resolution approaches, examine the models and algorithms of existing mining fleet management systems to identify its limitations, and propose a new system architecture that can optimize mine production and efficiency based on real-time data processing. In this article, we outline a discussion of real-time fleet management issues, a comparison of current fleet management tools, and an examination of potential system architectures.

Keywords FMS · Open-pit mine · Architectures · Multi-agent system · Real-time system

1 Introduction

Mining is generally considered as a basic industry, and it is characterized by difficult and dangerous working conditions, a heavy burden on the environment as well as a low level of high technology and automation [1]. A mining process (whether open or underground) is extremely complex, involve manual, physical, mechanical, and logistical operations with various interfaces and human decisions [2]. The mining

H. Bnouachir (✉) · H. Medromi
Engineering Research Laboratory (LRI), System Architecture Team (EAS) National and High School of Electricity and Mechanic (ENSEM), Hassan II University, Research Foundation for Development and Innovation in Science and Engineering, Casablanca, Morocco
e-mail: bnouachirhajar@gmail.com

M. Chergui
Computer Science and Smart Systems (C3S), National and High School of Electricity and Mechanic (ENSEM) Hassan II University, Casablanca, Morocco

M. Zegrari
Structural Engineering, Intelligent Systems and Electrical Energy Laboratory, Ecole Nationale Supérieure des Arts et Métiers, University HASSAN II, Casablanca, Morocco

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_52

industry is undergoing a profound transformation as a result of market volatility, the expansion of the cost base, and the evolution of global demand. In consequence, mining companies must adapt new operational and commercial models more efficiently than ever before. As a result, the Internet of Things, cloud mobility, and other digital tools and functions are being used by miners [3]. Digital transformation has a significant impact on how companies connect with their employees, communities, and governments in the mining sector. Furthermore, it benefits the environment at every level of the value chain, including mineral exploration and evaluation, remote sensing and satellite remote sensing, mining, ore processing, and production [4]. The digital transformation of the mining and metals sectors could generate a value of $320 billion over the next decade, with a potential of about $190 billion for the mining sector, and $130 billion for the metals' sector [4]. By visualizing data across the whole value chain at the beginning of the digital transformation process, mining companies can boost productivity, cut costs, and immediately improve production and security. In an integrated production process, they give stakeholders the knowledge they need to make better decisions. Additionally, mining companies can receive strategic information about their operations with the assistance of analytics and automatic learning algorithms. The performance, health, and qualities of a mineral can be predicted by miners by feeding these algorithms with real-time data and reviewing historical data. The combination of this information with a dynamic scheduling solution allows proactively [3]:

- Control mineral characteristics through strategies such as improving drilling and blasting and improving blending to achieve the required performance.
- Dynamically plan all mine operations based on predictive alerts.
- Improve asset health through asset predictive maintenance.
- Improve safety through fatigue monitoring or tracking of people and property.

New real-time data streams enabled by recent communications and information technology developments can be used to create new practical domains of implementation in the transportation industry throughout all manufacturing industries. In fact, a carriers [5] can, for example:

1. Find out about road traffic.
2. Be informed of which of its vehicles located.
3. Let clients know how their demands are progressing.
4. Maintain two-way communication with each vehicle in his fleet (transmission of information to drivers, confirmation of execution of a mission by the driver, etc.).

Real-time fleet management issues are included in the category of dynamic transport issues. We mention Powell's definition of a dynamic problem in order to refresh one's memory. Jaillet and Odoni [6]: "A problem is dynamic if decisions have to be made before all the information is known and modified when new information is refused". We will specify here the semantics of terms used throughout this document. Any service request that is given to the carrier and necessitates the use of at least one fleet vehicle is referred to as need. Any action that a vehicle must perform to accomplish one or more goals (such as moving, loading, unloading, repositioning) is referred to as a mission. We then refer to a cycle as a series of missions carried out or planned for a vehicle over a given time horizon. The problem of real-time fleet management consists of:

- Assigning a newly formulated requirement to a given cycle while previously established cycles are being executed.
- Ensuring that the cycles run smoothly when carrying out planned missions.

Fleet management system (FMS) generally refers to a broad range of solutions for the many fleet management problems connected to the vehicle fleet in the domains of logistics, distribution, and transportation. It involves focused planning, fleet operation supervision, and control in accordance with the transportation resources and application limitations that are now available. The FMS seeks to minimize costs while reducing risks, enhancing service quality, and enhancing a fleet's operational efficiency. Particularly, capital-intensive sectors with high operating costs are open-pit mining. About 50% of the operating costs of open-pit mines and even 60% of the operating costs of large open-pit mines are related to the transport and administration of materials. In open-pit mines, it has the greatest operational costs of any process control task [7, 8].

Utilizing real-time data, the fleet management system is being deployed into mines, especially open-pit mines, in order to increase mine production and efficiency. More specifically, the FMS aims to meet quality blending limits, feed the processing facility at the anticipated rate, and increase mining production while minimizing inventory management [9].

2 The Characteristics of Fleet Management Problems in Real Time

Based on the study carried out by Malca and Semet [5] in 2006, we can distinguish four groups of characteristics, namely: State Observation, Forecasting Anticipation, Evaluation, and Decision (Table 1).

Table 1 Characteristics of fleet management problems in real time [10]

State Observation	Time dimension	It is necessary that the dispatcher (or carrier) knows the geographical position of each vehicle at all times
	Updating information	The feedback from the fleet must be in real time in order to solve problems as quickly as possible
	Sizing of the fleet	The size of the fleet must be relative to any change in the identified need
Forecasting Anticipation	Management Horizon	**Short finite horizon:** Missions are determined in the near future based on the current objective
		Long finite horizon: Missions are determined on the basis of all known objectives
		Infinite horizon: Missions are determined on the basis of the execution of the missions of the previous objectives
	Future information	Future needs cannot be known with certainty, there are two types of problems to determine future events: **Stochastic problem:** Based onthe probability laws applied to the history of past events, it is possible to determine future objectives **Non-stochastic problem:** Future events are not taken into account through probabilistic laws
Evaluation	Short-term	It is requested to take into consideration the needs to be planned in the future in order to keep a certain flexibility in fleet management
	Calculation time	The impact of any changes or planning should be assessed as soon as possible
	Objective function	For the static case, the objective is to minimize costs and mission execution time, and for the dynamic case, the finality is not only to achieve all objectives but rather to maximize the quality of service provided to achieve them
Decision	Assignment and sequencing	The flexibility of the management system should make it easier to change objectives
	Acceptance period	The acceptance of any change in objective depends on the feasibility of such a change in the current context
	Queue	Establish a queue in case of several objectives that cannot be performed in parallel
	Nature of time constraints	Try to reduce the execution time of the assigned missions while respecting the quality of the desired service

3 Approaches to Solving Fleet Management Problems in Real Time

In a mining operation, a fleet management system is a decision- making tool that handles resources in an open-pit mine in real time. The FMS consults the database to get the necessary data concerning the operation's state and then takes appropriate action. When a new decision is needed, the FMS is recalled after the decisions have been applied in the operation [11]. Sending trucks by the shovel is one of the key decisions that the FMS will need to make. This choice must satisfy the demands for production while minimizing deviation and maximizing the output of active equipment, such as loaders and carriers. Although various heuristic strategies and approaches have been used since the 1960s to decide on truck distribution, none of them are capable of achieving all the planned approach at once [11]. Generally, fixed and dynamic models can be used to solve fleet management problems. We will focus on open-pit mine fleet management. As a result, two different types of assignments—fixed and dynamic—are included in the distribution of trucks to the open-pit mine.

3.1 *Fixed Distribution*

This approach puts a number of trucks on a fixed transport route (special excavator), where they will stay until the end of the shift. This route would not be changed until a shovel malfunctions or a significant event takes place. The fixed allocation technique is a static method, to put it another way. According to a number of factors, including the requirement for production, the availability of trucks in the fleet, etc., vehicles assigned to roads must operate on the same road during the shift. The various techniques used for the fixed assignment of trucks include the traditional methods based on the ideas of queue theory and programming. These techniques are all intended to establish the ideal ratio of trucks to shovels [12].

3.2 *Dynamic Distribution*

Based on the findings of this field's research, a fixed allocation strategy is not effective or useful for planning transport in huge mines [12]. With dynamic truck allocation, a certain activity shovel is given a specified number of trucks from the fleet at the start of the shift. However, after loading and unloading at the hoppers, these trucks will receive a new assignment of the distribution system each time, rather than serving a single shovel or road throughout the shift [9]. Several studies have shown that the flexible allocation strategy can significantly boost loading and transit capacity. Olson et al. [13] shown that the Bougainville copper mine's production increased by

13% by adopting the dynamic distribution of transport vehicles. Additionally, they showed increases of 10–15% in the productivity of the Barrick Goldstrike mine's gold production, 10% in the productivity of the LTV iron mine, and 10% in the production of coal from the Quintette mine. [13] According to a comparison study by Kolonja and Mutmansky [14] comparing the fixed allocation method and the variable allocation approach, the adoption of variable allocation significantly boosts mine operations' productivity.

4 Fleet Management Approaches to Open-Pit Mining: Truck Distribution

It has been shown that the majority of approaches to the distribution of truck loaders in open-pit mines can be classified into two categories: single-stage and multi-stage methods [8]. A single-stage FMS system finds the shortest paths between shippers and the destinations they are connected to, calculates the ideal number of truck trips needed for each open lane, and shovels the vehicles in one motion. Hauck (1973) created one of the earliest FMSs that is based on a one-step methodology. The majority of extraction FMSs, however, employ a multi-step process. This method involves resolving three sub-problems, with the results of each stage being utilized as an input for the subsequent sub-problem. These three related sub-problems are (a) shortest paths from all sources to all destinations, (b) best amount of material to produce for each path, and (c) real-time dispatching of trucks to excavators [9].

4.1 Criteria for the Distribution of Excavator Trucks in Open-Pit Mines

The ideal location for a transport truck is typically one that aims to optimize the satisfaction of one or more criteria, also known as distribution objectives. Transport trucks are assigned based on a variety of factors, that all aim to either directly or indirectly increase ore production or decrease machine inactivity [8].

Shortest path

One of the earliest descriptions of the operational problem in the literature on open-pit mining defines the shortest path as the shortest path between loading and the tipping points. To determine the shortest route between all of the loading and unloading points, Elbrond and Soumis [15] used a nonlinear programming network (NLP) algorithm. To choose the best path to connect the excavators to their destination, the majority of the linear programming (LP) algorithms created to date, including those presented by Temeng et al. [16] and Temeng et al. [17], use the Dijkstra algorithm to find the shortest path between the source and the well.

Optimization of production and allocation of trucks

The second sub-problem is referred to as production optimization. Systems for production optimization have already been accomplished utilizing complete Mixed Linear Programming (MILP), Linear Programming (LP), and Nonlinear Programming (NLP). The use of the aforementioned techniques results in either the total amount of tonnage to be delivered from the mine's various loading areas to the destinations or the number of truck trips required to fulfill each trajectory's target production rate within a certain period [9].

Real-time distribution

In a fixed truck distribution mine, real-time decision-making on the destination of trucks was initially used in the early 1960s using radio communication technologies to connect the dispatcher and truck drivers. However, based on the usage of modern computers, real-time fleet management in mining systems is divided into three major classes [8]:

- Systems with fixed or locked allocations.
- Semi-automated systems.
- Fully automated systems.

5 Bucket Truck Distribution Strategies

The objective of optimizing a distribution system is to optimize productivity. The distribution methods considered in the literature are based in part on reducing the waiting time for trucks to wait for shovels to be used. Therefore, if the time lost in the queue is reduced, truck use will increase. The fleet management problems presented above are based on the concept of real-time management, in our case the distribution of trucks in time [8].

5.1 The Approach of One Truck for N Pellets

The most typical tactic employed in mining operations is the one truck for n pellets' method (heuristics). An excavator could be sent in place of a truck in response to a truck operator's request for a different assignment. The dispatcher's judgment or logical operating procedure, which typically employs one of the heuristic approaches listed below, determines which excavator the truck is allocated to. The shovel with the most potential receives the truck. In general, a one-step approach is used to accomplish this strategy.

5.2 The M Methods—Trucks-For-One-Excavator

The “m method” (trucks-for-one-shovel) is focused on a multi-step strategy; choices regarding truck allocation will be made as the trucks are being sent, taking into account one shovel at a time. To be more explicit, the shovels are initially prioritized based on how late they are in respect to the manufacturing schedule. The dispatcher then gives the shovel that is highest on the priority list of the best truck.

5.3 The M-Trucks to the Act N Methods—Shovel

Based on the anticipated availability of trucks and excavators, the dispatcher assigns the asking truck to the optimal shovel while simultaneously evaluating the m incoming trucks and n available shovels. This process for matching m trucks to n shovels is multi-step in nature. The requester’s truck is the only one that gets affected. In this method, M must be more than or equal to n.

6 Distribution Algorithm Limitations

The development of fleet management systems for use in surface mines has attracted a lot of research attention [9]. The currently chosen algorithms and models still have a lot of shortcomings and limitations. The models’ poor connection to strategic plans, particularly the short-term strategy, is a significant flaw. Because the deposit is typically divided into big polygons by strategic plans, there is a weak link in the chain. The algorithms’ ability to ignore operational and geological variables that have an impact on fleet management systems is another problem. Other aspects that almost all fleet management systems have so far disregarded include production losses brought on by big equipment movements, fleet heterogeneity, and shortest path dynamics, particularly in large mining operations:

- Link between strategic and operational level plans.
- Accounting for uncertainty.
- Mobility and access to equipment.
- Dynamic determination of the best trajectory.
- Dispatching in real time based on the transshipment issue.

Current algorithms still have limitations [9], namely:

- The operational and strategic parts of the plan are not connected; thus, the proposed fleet management systems do not allow for the execution of both short-term and long-term objectives.
- The majority of models are deterministic, assume a fixed average grade for each size front, and do not take into consideration the mine’s whole life.

- One of the main causes of the production rate variation is that the current systems do not account for the tons lost when moving shovels from one level to another.
- By completely modeling an open-pit mining operation and accounting for the heterogeneity of the vehicles, the proposed models must be as realistic as feasible.
- Dynamically calculating the shortest paths by taking into consideration the truck's present location, its intended destination, and the amount of time needed to get there given any traffic jams that may be present.
- It is advised to use a transshipment problem strategy in the dispatching operation rather than a transport or assignment technique. There are delivery and demand points as well as transshipment points in this kind of issue where commodities are moved from suppliers to demand locations. System stocks or network intersection nodes may be regarded as transshipment locations in the mining industry.

Mine fleet management solutions are offered by numerous companies worldwide. A preliminary evaluation is proposed in [18] of some industry systems using the ISA 95 criteria.

7 What Architecture to Choose?

With the advent of automation, computer technologies, and industrial software at all levels of production systems, we now have increasingly powerful means at our disposal to transfer information from the field to information systems. Senehi and Kramer define the piloting architecture as a description of the composition and structure of a piloting system [19], and according to the definition of Trentesaux in 2009 [20], piloting architectures can be composed of biological or artificial beings with a certain "intelligence" and communication or interaction capacities. For Trentesaux, his beings correspond to "entities". Management architecture therefore includes all the entities of the system and the relationships between them. In this architecture, level i entities participate in the management of their level and are subjected to level $i + 1$ management. As seen previously with Trentesaux, this term can refer to biological beings (human operator, production manager, etc.) or artificial beings. When they are artificial, they can be modeled by the Multi-Agent, [21], Holonique [22], Actor [23], Fractal [24], or Modelon [25] models, for example. In our case, we are interested in the design and implementation of a real-time mining fleet management information system. Fleet management refers to the applications, tools, technologies, and practices that help companies optimize the use of their work vehicles from a central platform:

- Database information software.
- GPS telematics and tracking software.

However, the question that arises is: what optimization approach should be adopted to ensure efficiency and better service quality? The system must be able to access information systems at all operational levels in real time and automatically adapt to their needs. Then, it is necessary to be able to apply complex route search and optimization algorithms while managing the hot integration of new IS. To do this, it is necessary to think of architecture that facilitates the communication of the system with its environment and even the communication between the different entities of the system itself. In addition, the system must not be affected and must be reliable in the event of disturbances. Indeed, it must allow automatic registration and unsubscription of SIADs without having to modify the source code. Without forgetting, of course, the complex calculation, it must perform to find truck-shovel assignment solutions and calculate the best solutions in terms of three criteria: execution times (response time, processing time, time to acquire real data), cost, and the facility for operator-system interaction. To set up such a system, another question arises: which technology or methodology to choose? In the field of software engineering, new structural and architectural concepts have emerged as a result of the growing demand for increasingly complex information systems and software. The most common architectures are component, service, and agent-based architectures.

7.1 Component-Based Architecture

Component-based technology is an architecture extracted from object-oriented approaches [26], and based on the concept of the electrical circuit, it consists of considering computer software as a set of several components. According to Chardigny [26], "a component is a software element that is composable without modification, can be distributed independently, encapsulates a functionality, and adheres to a component model". It has input and output interfaces that allow it to interact with other components. Component-based programming finds its effectiveness through reuse. The notion of component fills the limits of the object approach in terms of granularity of reuse, thus ensuring a better safety of component reuse [27]. Indeed, unlike an object that may make unsecured reuses (following calls to external services without explicit specification), a component can only use well-specified services given its design and details on its interfaces.

7.2 Services-Oriented Architecture

Service-Oriented Architecture (SOA) is a new application organization that allows interaction between remote components of the application through services. Each service is accessible through standards and message exchange protocols Rouillard et al. [28]. The notion of service can be represented here by an object produced by a supplier and consumed by a customer. SOA proposes a new concept that facilitates the exchange of messages, is reusable, and has good security (use of standardized protocols). The service is not exclusively a web service, but perhaps any type of

Table 2 Architectures' comparison

Entities	Component-based architecture	Service- Oriented Architecture	Agent-based architecture [31, 32]	
Autonomy	A component can be distributed autonomously	Service autonomy	Holon-based [33]	Multi-agent
			A Holon is an autonomous entity	An agent is an autonomous entity
Cooperation	It has interface that allows it to interact with other components	The organization of isolated software applications through the implementation of an infrastructure	The process by which a group of Holons creates and implements mutually agreeable goals	Complex problems can be solved with the combined knowledge of several agents
Recursivity	–	–	The Holon notion is recursive, which distinguishes the Holonic technique from the multi-agent method	
Direct integration	–	–	Representation of physical elements by agents or Holons	

service that respects one or more protocols and a precise description made available to the customer, for example, for Web Services Description Language (WSDL).

7.3 Agent-Based Architecture

An agent is an autonomous entity capable of communication, with private knowledge and behavior and its own capacity for execution [29]. The combination of the knowledge of several agents and their cooperation makes it possible to solve complex problems and to give more impetus to resolution skills [30].

7.4 Architecture Comparison

We will present in Table 2 a comparison between the types presented above.

8 Perspective and Future Work

Currently, a variety of domains use fleet management systems (FMS) to coordinate traffic and logistics services [34]. However, in open, dynamic, and developing contexts where flexibility and autonomy are required and desired, their conventional and traditional control architecture becomes a serious impediment [35]. In this regard, we have tried to understand the real-time fleet management problems, their characteristics, and proceeded approaches, examine the models and algorithms of existing mining FMS with the aim to identify their limitations in order to propose an intelligent distributed FMS architecture [36, 37] for an open-pit mine. These latter enable real-time control and decision-making of mining vehicles allowing the FMS to improve their agility and response. Our architecture presents many contributions to the field that allow the FMS to meet the interoperability and autonomy requirements of the most widely used standards in the field, such as ISA 95.

9 Conclusion

To address fleet management issues in open-pit mines, researchers have created and used a variety of algorithms. In front of our review, we will develop the dispatching algorithms, detail and implement each element of the architecture for the smart fleet management system.

Acknowledgements Mohammed VI Polytechnic University of Benguerir, in Morocco, provided support for this study. We are grateful to our university colleagues who offered their knowledge and skills, which considerably aided the research.

References

1. Seoble M (1995) 957353 Canadian mining automation evolution: the digital mine en mute to minewtde automation. 1
2. Uronen P, Matikainen R (1995) The intelligent mine. IFAC Proc 28:9–19. https://doi.org/10.1016/S1474-6670(17)46739-5
3. Accenture digital mining connecting the mine from pit to port, from sensor to boardroom, for improved safety and productivity
4. Pan Pacific Perth, Australia (2019) Digital mines: building fully autonomous mines from pit to port
5. Frédéric Semet FM (2006) Les problemes de gestion de flotte en temps reel. INFOR: Inf Syst Oper Res 44:299–330. https://doi.org/10.1080/03155986.2006.11732754
6. Powell WB, Jaillet P, Odoni A (1995) Chapter 3 Stochastic and dynamic networks and routing. In: Handbooks in operations research and management science. Elsevier, pp 141–295
7. Curry JA, Ismay MJL, Jameson GJ (2014) Mine operating costs and the potential impacts of energy and grinding. Miner Eng 56:70–80. https://doi.org/10.1016/j.mineng.2013.10.020
8. Alarie S, Gamache M (2002) Overview of Solution strategies used in truck dispatching systems for open pit mines. Int J Surf Min Reclam Environ 16:59–76. https://doi.org/10.1076/ijsm.16.1.59.3408
9. Afrapoli AM, Askari-Nasab H (2019) Mining fleet management systems: a review of models and algorithms. Int J Min Reclam Environ 33:42–60. https://doi.org/10.1080/17480930.2017.1336607
10. Gafert M (2021) Challenges for future automated logistics fleet interactions. 6

11. Moradi Afrapoli A, Tabesh M, Askari-Nasab H (2019) A multiple objective transportation problem approach to dynamic truck dispatching in surface mines. Eur J Oper Res 276:331–342. https://doi.org/10.1016/j.ejor.2019.01.008
12. Ahangaran DK, Yasrebi AB, Wetherelt A, Foster P (2012) Real –time dispatching modelling for trucks with different capacities in open pit mines / Modelowanie w czasie rzeczywistym przewozów ciężarówek o różnej ładowności w kopalni odkrywkowej. Arch Min Sci 57:39–52. https://doi.org/10.2478/v10267-012-0003-8
13. Zhang L, Xia X (2015) An integer programming approach for truck-shovel dispatching problem in open-pit mines. Energy Procedia 75:1779–1784. https://doi.org/10.1016/j.egypro.2015.07.469
14. Kolonja B, Kalasky DR, Mutmansky JM (1993) Optimization of dispatching criteria for open-pit truck haulage system design using multiple comparisons with the best and common random numbiers. 9
15. Elbrond J, Soumis F (1987) Towards integrated production planning and truck dispatching in open pit mines. Int J Surf Min Reclam Environ 1:1–6. https://doi.org/10.1080/09208118708944095
16. Temeng VA, Otuonye FO, Frendewey JO (1997) Real-time truck dispatching using a transportation algorithm. Int J Surf Min Reclam Environ 11:203–207. https://doi.org/10.1080/09208119708944093
17. Temeng VA, Otuonye FO, Frendewey JO (1998) A Nonpreemptive Goal Programming Approach to Truck Dispatching in Open Pit Mines
18. Bnouachir H, Chergui M, Zegrari M et al (2022) Smart fleet management system based on multi-agent systems: mining context. In: Kacprzyk J, Balas VE, Ezziyyani M (eds) Advanced intelligent systems for sustainable development (AI2SD'2020). Springer International Publishing, Cham, pp 748–761
19. Senehi MK, Kramer TR (1998) A framework for control architectures. Int J Comput Integr Manuf 11:347–363. https://doi.org/10.1080/095119298130688
20. Trentesaux D (2009) Distributed control of production systems. Eng Appl Artif Intell 22:971–978. https://doi.org/10.1016/j.engappai.2009.05.001
21. Chergui M, Chakir A, Medromi H (2019) Smart IT governance, risk and compliance semantic model: business driven architecture. In: 2019 Third world conference on smart trends in systems security and sustainability (WorldS4). IEEE, London, United Kingdom, pp 297–301
22. Arthur Koestler (1967) The ghost in the machine. 404
23. Mbobi M, Boulanger F (2006) Le paradigme acteur dans la modelisation des systemes embarques. In: 2006 Canadian conference on electrical and computer engineering. IEEE, Ottawa, ON, Canada, pp 418–421
24. Ryu K, Jung M (2003) Agent-based fractal architecture and modelling for developing distributed manufacturing systems. Int J Prod Res 41:4233–4255
25. Ueda K (1992) A concept for bionic manufacturing systems based on DNA-type information. In: Human aspects in computer integrated manufacturing. Elsevier, pp 853–863
26. Chardigny S, Seriai A, Oussalah M, Tamzalit D Extraction d'Architecture à Base de Composants d'un Système Orienté Objet. 16
27. Meijler TD, Nierstrasz O, Beyond objects: components. 26
28. Rouillard J, Vantroys T, Chevrin V (2007) Les architectures orientées service. Une approche pragmatique des SOA
29. Chergui M (2017) Conception et Réalisation d'une Plateforme de Gouvernance des Systèmes d'Information à base des workflow inter-organisations, du Web sémantique et des Systèmes Multi-agent. Université Hassan II, Casablanca. Ecole Nationale Supérieure d'Électricité
30. Leriche S (2006) Architectures à composants et agents pour la conception d'applications réparties adaptables. PhD Thesis
31. Najjari H, Seitz M, Trunzer E, Vogel-Heuser B (2021) Cyber-physical production systems for SMEs-A generic multi agent based architecture and case study. In: 2021 4th IEEE international conference on industrial cyber-physical systems (ICPS), pp 625–630

32. Noureddine DB, Krichen M, Mechti S et al (2021) An agent-based architecture using deep reinforcement learning for the intelligent internet of things applications. In: Saeed F, Al-Hadhrami T, Mohammed F, Mohammed E (eds) Advances on smart and soft computing. Springer, Singapore, pp 273–283
33. Moise G, Moise P-G, Moise P-S (2018) Toward holons-based architecture for medical systems. In: 2018 IEEE/ACM international workshop on software engineering in healthcare systems (SEHS), pp 26–29
34. Barnewold L, Lottermoser BG (2020) Identification of digital technologies and digitalisation trends in the mining industry. Int J Min Sci Technol 30:747–757. https://doi.org/10.1016/j.ijmst.2020.07.003
35. Zhang S, Lu C, Jiang S et al (2020) An Unmanned intelligent transportation scheduling system for open-pit mine vehicles based on 5G and big data. IEEE Access 8:135524–135539. https://doi.org/10.1109/ACCESS.2020.3011109
36. Bnouachir H, Chergui M, Machkour N et al (2020) Intelligent fleet management system for open pit mine. Int J Adv Comput Sci Appl 11:6
37. Benlaajili S, Moutaouakkil F, Chebak A et al (2020) Optimization of truck-shovel allocation problem in open-pit mines. In: Hamlich M, Bellatreche L, Mondal A, Ordonez C (eds) Smart applications and data analysis. Springer International Publishing, Cham, pp 243–255

An Epidemiological SIS Malware Spreading Model Based on Markov Chains for IoT Networks

J. Flórez, G. A. Montoya, and C. Lozano-Garzón

Abstract IoT technology has been on an uprising these last years, and the number of devices connected to the Internet is likely to keep increasing. As such, the amount of attacks targeting these devices is at an all-time high, and a big percentage of all cyberattacks are focused on IoT devices. This creates the necessity of proposing models to estimate the impact of malware on an IoT network that helps the proposal of countermeasures to protect the network and reduce the possible costs of an attack. In this sense, we propose a stochastic epidemiological SIS model to analyze the behavior of an interconnected network of IoT devices that have been infected by malware. To fulfill this goal, we formulated the initial SIS model, then, implemented this model using Markov Chains, and validated our model by comparing it to the Gillespie simulation algorithm.

Keywords Epidemiological models · Markov chains · Internet of things · Stochastic models · IoT security

1 Introduction

The Internet of things is defined as the network of physical objects with sensors, software, and other technologies to gather information and share it with other devices through the Internet. As we can see on [1], IoT devices are used in a wide range of innovative applications; from a smart environment that can predict natural disasters and communicate with each other; to a smart home that can help people remotely control home appliances depending on their needs, and many more examples such

J. Flórez (✉) · G. A. Montoya · C. Lozano-Garzón
Universidad de los Andes, Bogotá, Colombia
e-mail: jd.florezg1@uniandes.edu.co

G. A. Montoya
e-mail: ga.montoya44@uniandes.edu.co

C. Lozano-Garzón
e-mail: calozanog@uniandes.edu.co

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_53

as smart hospitals, smart agriculture, and smart retailing. This type of networks are highly vulnerable to different types of attacks such as viruses used to gain access to devices or software, without user knowledge, to perform malicious tasks; worms that duplicate into thousands of copies, allowing them to alter operations; and spyware used to collect user information without its knowledge [2].

We will be focused on botnet attacks, which is a type of malware that infects an IoT device and starts to spread to the rest of the devices on the network. These Bowith sensors, software, and other technologies to gather information and sharemple are [3], which corresponds to the Mirai botnet. This one was the first botnet that infected nearly 65,000 IoT devices in its first 20 h and got to a population of 300,000 infections. As we can see, botnets have become a wide world phenomenon, and there have been detected a considerable amount of cases [4]. In this sense, with the increase of IoT technologies and the amount of malware designed to attack these devices, we start to need a method for analyzing the behavior of malware and how it spreads on a network of interconnected devices. To fulfill this objective, we propose an epidemiological model based on networks that simulate the spreading of malware.

This paper is organized as follows: Sect. 2 describes what mechanisms have been used to model the spread of malware in a network of connected IoT devices. In Sect. 3, we propose our approach based on a stochastic SIS model and the metrics used to measure the spread of malware on a network. In addition, we will discuss the details of the implementation of the proposed model and we show how the algorithms work and how they compare to a simulation approach and a real scenario using the Mirai malware in Sect. 4; finally, in Sect. 5, we present the conclusions.

2 Related Works

To analyze the impact of malware on an interconnected network of devices, there have been some proposals on how to model the evolution of the malware over time. As seen on [5], there are some techniques used to detect when a device has been infected with malware and when a file could be dangerous to the devices on the network. These approaches include the use of machine learning to analyze different files and find a pattern regarding the dangerous files; that is, analyzing the network behaviors to understand the changes on the network when it has been infected with a virus and the study of the control graphs of files that may have malware to differentiate them with non-threatening files. The main difference is that this paper is centered around techniques to predict the impact that malware might have on the network instead of trying to detect when malware has infected our devices.

On the side of epidemiological models centered around networks, there are several to model different behaviors of the virus spreading through the network. The main difference is the initial hypothesis that we hold for each model. The most usual epidemiological models on networks can be found on [6] and [7] and are primarily the SIS, SIR, SIRS, and SEIRS models. The use of these epidemiological models on IoT networks has been discussed before in papers such as [8] where they propose an

SEIRS model to describe the spread of IoT worms such as Mirai and the vulnerability that they create for the Internet. This work uses a deterministic SEIRS model approach to conclude to mitigate the frequency of IoT botnet attacks with improved user information.

In [9], the IoT SIS, SIR, and SEIR deterministic models are discussed to predict the spread of malware. They propose a stochastic SIRS model and present an analytical conclusion and some simulations of the model.

The goal of this paper consists of proposing one of the pure stochastic SIS models, using it on an IoT network. This allows us the analysis of specific IoT networks, the characterization of the most vulnerable nodes, and the calculation of the possible costs at the time of an infection, which is in contrast to the models mentioned that focus on the analysis of the SIS model for a general network, not for any specific topology. In addition, the model will be tested by comparing it to the Gillespie SIS simulation algorithm and using a specific network with the Mirai malware data.

3 Stochastic SIS Model on Heterogeneous Networks

3.1 Definition of the Model

To model the spread of the malware through the network of IoT devices, we are going to use the SIS epidemiological stochastic model. This model has the following components: $N = \{1, ..., n\}$ is set of all devices of the network; τ and γ correspond to the infection and recovery rate, respectively; and A corresponds to the adjacency matrix of the network. Using these parameters of the network, we define a continuous time Markov chain as follows: $X(t)$ is a random variable that corresponds to the number of infected nodes at time t; the set of states of our Markov Chain will be the set 2^N or the set of subsets of N; and the generator matrix of the Markov chain U is described in the following equation:

$$U_{(i_1,i_2,...,i_n)\to(j_1,j_2,...,j_n)} = \begin{cases} \tau * \sum_{p=1}^{n} A_{p,k} * i_p & \text{if } j_k = i_k + 1 \\ \gamma & \text{if } j_k = i_k - 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Since this is a generator matrix of a Markov chain, the diagonal entries of the matrix U will be the negative sum of the elements of the corresponding row.

3.2 Metrics

Using this model, we can gather information on the network and get a general idea of how the network would behave in the case of a malware attack. We will use the usual measures that can be defined on Markov chains found in [10]:

- Expected time until the infection ends: This time is defined as the expected time before we reach the absorbent state; in our model, this would be the (0, 0, 0, ..., 0) state which is the state where every node is susceptible, and the malware is no longer on the network. In a continuous time Markov chain, the expected time before we reach the absorbent state $(j_1, j_2, ..., j_n)$ from the initial state $(i_1, i_2, ..., i_n)$ is calculated by solving:

$$\begin{cases} t_{(j_1,j_2,...,j_n)} = 0 & \text{if } j_k = i_k \forall k = 1, ..., n \\ -\sum_{i=1}^{n} U_{(i_1,i_2,...,i_n)\to(j_1,j_2,...,j_n)} * t_{(j_1,j_2,...,j_n)} = 1 & \text{otherwise} \end{cases} \tag{2}$$

- Number of expected nodes infected on time t: Let us take Q as the transition matrix derived from the generator matrix U of the Markov chain, and E(i) the number of nodes infected on the state i, i.e., $E(i_1, ..., i_n) = \sum_{k=1}^{n} i_k$, then the expected number of nodes infected on the time t is given by:

$$\sum_{j=1}^{n} Q_{i,j}^{t} * E(i) \tag{3}$$

- Critical node: We will define the critical node as the node where if the infection starts, the time expected until it ends will be longer. In other words, if for each node $n \in N$, we calculate the expected time until the infection ends given by $t_{(0,0,...,n,...,0)}$, then the critical node will be the node where this value is the highest.

In terms of algorithmic complexity, if our network contains n devices, then we can see by our definition of the Markov chain that we will have $G = 2^n$ states, and that the generator matrix of the Markov chain will be of size $O(G^2)$, which will determine the space complexity of this algorithm. The time complexity can be divided according to each of the steps defined on the previous section. Firstly, building the Markov chain has time complexity of $O(G^2)$ since we have to compute every entry of the generator Matrix. Secondly, calculating the expected time until the infection end has complexity $O(G^3)$ since we have to solve a linear system of G equations. Finding the critical node is also of the same order since it calculates the expected time before the infection ends and takes the highest. Finally, calculating the number of expected nodes infected on time t has a time complexity of $O(t * G^3)$ since we have to multiply the matrix Q with itself t times.

3.3 Pseudocode

The model implemented can be summarized in four smaller algorithms:

- Algorithm 1 allows us to build the continuous time Markov chain given the adjacency matrix A, the infection rate τ, and the recovery rate γ:
  Where $len(A)$ is the number of rows of the matrix A; $0_{2^n,2^n}$ is the matrix with only 0 of size 2^n; $toBinary(i)$ returns an array of the digits of i in binary: and $sum(a)$ is the sum of all elements of the array a.

Algorithm 1 Creating the Markov Chain

```
procedure CREATEMARKOVCHAIN(A,γ,τ)
    n ← len(GeneratorMatrix)
    GeneratorMatrix ← 0_{2^n,2^n}
    for i ∈ {0, 1, ..., 2^n} do
        for j ∈ {0, 1, ..., 2^n)} do
            a ← toBinary(i).
            b ← toBinary(j).
            if sum(a) = sum(b) + 1 then
                GeneratorMatrix[i][j] ← τ.
            end if
            if a = [a_1, ..., a_k, ...], b = [a_1, ..., a_k − 1, ...] then
                GeneratorMatrix[i][j] ← γ Σ_{p=0}^{n} A[p][k] ∗ b[p]
            end if
            close;
        end for
        close;
    end for
    for i ∈ {0, 1, ..., 2^n} do
        GeneratorMatrix[i][i] ← −sum(GeneratorMatrix[i])
    end for
    close;
    return GeneratorMatrix
end procedure
```

Algorithm 2 Expected time before the infection ends

```
procedure STEADYTIME(GENERATORMATRIX)
    t ← []
    n ← len(GeneratorMatrix)
    for i ∈ {0, 1, ..., n} do
        if 2^i < n then
            t.push(ExpectedTime(GeneratorMatrix, 2^i, 1)).
        end if
        close;
    end for
    return t
end procedure
```

- Algorithm 2 calculates the expected time before the infection ends: Where $t.push(a)$ appends the element to the end of the array t, and the function ExpectedTime (CMCT, initialnode, endnode) is a function that, given the Markov Chain, the initial node and the end node, returns the expected time to get from the initial node to the final node. The function $ExpectedTime$ was taken from [10].
- Algorithm 3 gets the expected number of nodes infected after n instants given an initial infected state.

Algorithm 3 Expected infections

```
procedure EXPECTEDINFECTIONS(GENERATORMATRIX, T, INITIALSTATE)
    ProbNSteps ← GeneratorMatrix^t
    initialVector ← 0_{1,len(GeneratorMatrix)}
    initialVector[InitialState] ← 1
    probabilities ← initialVector ∗ ProbNSteps
    n ← len(GeneratorMatrix)
    expectedValue ← Σ_{i=1}^{n} probabilities[i] ∗ toBinary(i)
    return expectedValue
end procedure
```

– Algorithm 4 gets the critical node of the network:

Algorithm 4 Critical node of the network

```
procedure CRITICALNODE(GENERATORMATRIX)
    nodeValues ← SteadyTime(GeneratorMatrix)
    max ← 0
    i ← 0
    maxValue ← nodeValues[0]
    for val ∈ nodeValues do
        if val > maxValue then
            maxValue ← val
            max ← i
        end if
        close;
        i ← i + 1
    end for
    close;
    return max
end procedure
```

Here, the $SteadyTime(GeneratorMatrix)$ function corresponds to the algorithm 2.

4 Implementation and results

4.1 Implementation

The algorithms were written in R version 4.0.3 and the packages “shiny”, “shinydashboard”, and “igraph” were used to show the results. Also, we used the package “markovchain” [10] to handle the continuous time Markov chain. The code for the stochastic model can be found at https://github.com/Enguenye/ModeloSISRedes.

All the tests were run on a Windows 10 computer with 32 GB of RAM and an Intel(R) Core(TM) i5-8600K CPU @ 3.60GHz.

4.2 Comparison with the Gillespie Algorithm

To validate our model, we will compare it against the Gillespie algorithm on SIS networks. This algorithm was taken from [11] and is a simulation of the Markovian process. We will compare our model with this simulation algorithm as follows:

- First, we will define the following parameters for both models:
  - τ, γ : the infection and recovery rates. Let us recall that the infection rate denotes how often a node can infect another node in the network, and the recovery rate tells us how often a node recovers after being infected.

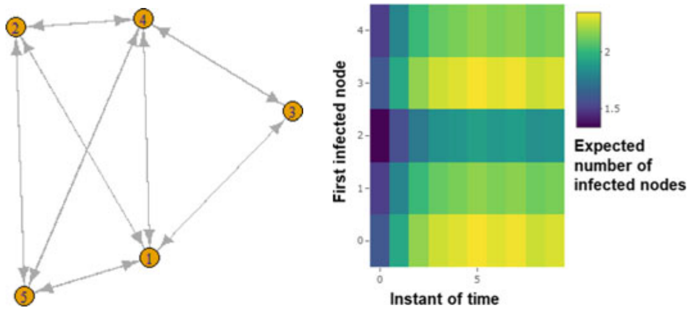
  - A: The adjacency matrix of the network.
  - n: Number of samples of the Gillespie algorithm, i.e., the number of times the simulation will be run.
  - t: Maximum time steps to calculate and compare.

- Now, we will run our model and the Gillespie algorithm changing the parameters to some values of the following set $\{n, t, A : n = 10, 20, 30, 40, 50, 60 \wedge t = 8, 11, 14, 17, 20 \wedge 5 \leq |A| \leq 12\}$.
- We should now get the results from our model. Let us consider the following network: Fig. 1a shows our network scenario, which is composed of 5 nodes and the connections between them. Now, we run our Markov chain-based model and get the following results:

  - The heat map 1b represents the evolution in time of the expected number of infected nodes depending on the first infected node of the network. As we can see, the malware spreads to a higher amount of nodes if it starts on the node 1 or 4 since they both are connected to every other node on the network. In addition, if the malware starts infecting the node 2, the spread is significantly less since that node is the most isolated. On the other hand, we run the Gillespie algorithm for the same network with $n = 30$ samples.
  - Figure 2a also represents the evolution in time of the expected number of infected nodes depending on the first infected node of the network but given by the Gillespie algorithm. We can see that both heat maps are very similar, and this one also shows that node 1 and 4 are the most impactful, and node 3 is the least impactful.
  - If we compare both distributions with the *t-student* test, we get a $p - value$ equal to $3.65475584620862e - 12$, which means our model is successfully validated in comparison with the Gillespie algorithm. If we repeat this process with all the generated scenarios, we end up with a list of p-values with a distribution given in Fig. 2. This process showed the following properties:

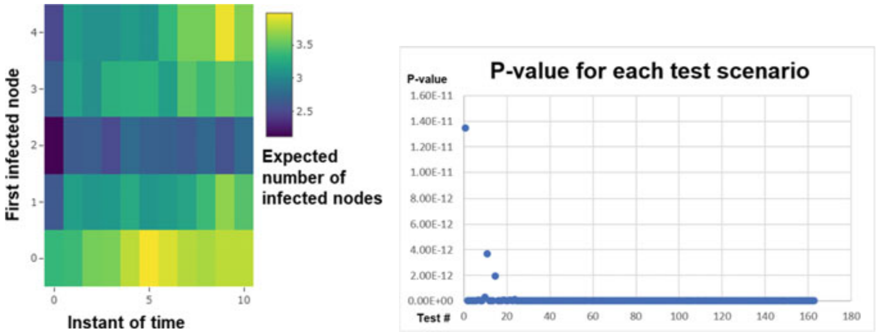
    Minimum: $1.05 * 10^{-54}$
    Maximum: $1.34 * 10^{-11}$
    Average: $1.19 * 10^{-13}$
    Standard deviation: $1.09 * 10^{-12}$. These low values, as we said before, confirm the successful validation of our model in comparison with the Gillespie algorithm.

Figure 2b shows that the distributions given by our model and the Gillespie model are close since the p-values of every comparison are infimum. This means that we can confidently say that the Markov model is suitable and correct to track the number of infected individuals in the network. Also, we note that the first results for our p-value tend to be higher, since the Gillespie algorithm uses less samples in earlier scenarios which means that the approximation given by the simulation algorithm is not as good.



(a) Network scenario (b) Heat map of the Markov algorithm

Fig. 1 Network scenario and heat map of the Markov algorithm



(a) Heat map of the simulation (b) Comparison of the results

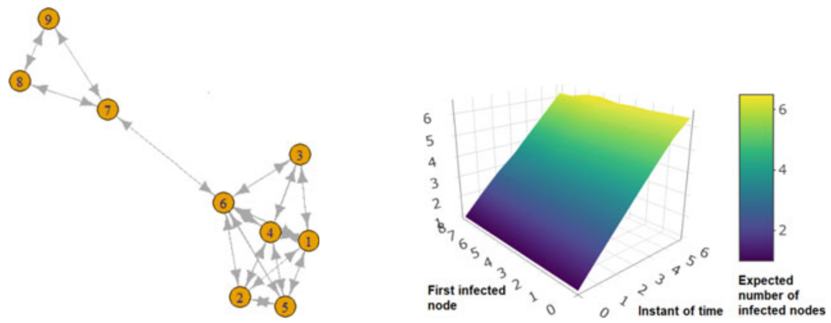
Fig. 2 Heat map of the simulation and comparison of the results

4.3 Proof of Concept Scenario

We will see how our proposal works with a real case using the Mirai malware as a basis and the network shown in Fig. 3a. This network consists of two subgraphs, one which has the nodes 7, 8, 9 and one with the nodes 1, 2, 3, 4, 5, 6 that are conected via nodes 6 and 7.

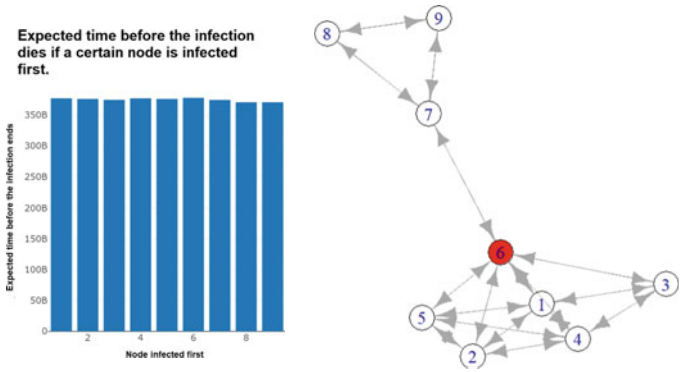
According to [3] as the reference for the Mirai spreading and recuperation rates, we select $\tau = 1.8$, $\gamma = 0.1$ as the infection and recovery rates, respectively. Using our Markov chain algorithm with these parameters and the graph adjacency matrix, we calculate how the malware spreads through the network (Figs. 3b and 4a).

We can see in Fig. 3b that the spread in nodes 7, 8, and 9 is lower than the other nodes since they are more isolated and that node 6 has the most impact regarding number of infected nodes. Even though we can draw some conclusions from this graph, since the spreading rate is much higher than the recuperation rate, the infections tend to increase rapidly, and it is difficult to discern which node contributes



(a) Test network (b) Expected number of infected nodes on time t

Fig. 3 Test network and evolution of the Malware spreading



(a) Expected time before the infection ending (b) Critical node of the network

Fig. 4 Test network and evolution of the Malware spreading

more to the spreading of the Malware. To get a better result, we calculate the expected time before the disease ends if a node was infected first (Fig. 4a).

From Fig. 4a, we can see that the highest time before the infection ends corresponds to the 6th node, which is also shown in the network graph in Fig. 4b. Also, since the infection rate is much higher than the recovery rate we see that i would take a long time for the infection to die in the described network.

In Fig. 4b, we can see how this algorithm could be used to detect the most vulnerable nodes in a network for Mirai malware. In our particular case, it would make sense that node 6 is the most critical node on the network since this node connects the two subnets. We then could use monitoring or other protective tools prioritizing node 6 since it is the most important node regarding malware spread.

5 Conclusions

A stochastic epidemiological SIS model based on Markov chains was proposed to analyze the behavior of malware spreading in an IoT network. By obtaining very low p-values between our model and the Gillespie algorithm, our model was successfully validated. As a result, in this work, a stochastic epidemiological SIS model based on Markov chains is suitable for analyzing the malware spreading in an IoT network, in this case, by analyzing the well-known Mirai malware in some proof of concept network scenarios. In this sense, our model could be used for tracking the number of infected individuals in the network to identify critical nodes and, then, take the best network decisions to stop the malware from spreading. In other words, and according to our study, by calculating the expected time before the disease ends, it was possible to identify the nodes that most contributed to the malware spreading in the network. This finding could be very useful for IoT network administrators to improve their decision-making in terms of network security.

References

1. Farooq M, Waseem M, Mazhar S, Khairi A, Kamal T (2015) A review on Internet of Things (IoT). Int J Comp Appl. https://doi.org/10.5120/19787-1571
2. Karanja M, Masupe S, Jeffrey M (2017) Internet of Things Malware : a survey. Int J Comp Sci Eng Surv. https://doi.org/10.5121/ijcses.2017.8301
3. Antonakakis M, April T, Bailey M, Bernhard M, Bursztein E, Cochran J, Durumeric Z, Alex Halderman J, Invernizzi L, Kallitsis M, Kumar D, Lever C, Ma Z, Mason J, Menscher D, Seaman C, Sullivan N, Thomas K, Zhou Y (2017) Understanding the Mirai Botnet. In: 26th USENIX security symposium (USENIX Security 17), USENIX Association. ISBN 978-1-931971-40-9
4. Alieyan K, Almomani A, Abdullah R, Almutairi B, Alauthman M (2020) Botnet and Internet of Things (IoTs): a definition, taxonomy, challenges, and future directions. In: Security, privacy, and forensics issues in big data, pp 304–316
5. Ngo QD, Nguyen HT, Le VH, Nguyen DH (2020) A survey of IoT malware and detection methods based on static features. ICT Express. https://doi.org/10.1016/j.icte.2020.04.005
6. Kiss IZ, Miller JC, Simon PL (2017) Mathematics of epidemics on networks. From Exact to Approximate Models. Springer
7. Del Rey Angel M, Acarali D, Rajarajan M, Komninos N, Zarpelão BB (2019) Modelling the spread of Botnet Malware in IoT-based wireless sensor networks. Secur Commun Netw. https://doi.org/10.1155/2019/3745619
8. Gardner MT, Beard CC, Deep M (2017) Using SEIRS epidemic models for IoT Botnets attacks. In: 13th International conference of the DRCN 2017—design of reliable communication networks
9. Mahboubi A, Camtepe S, Ansari K (2020) Stochastic modeling of IoT Botnetspread: a short survey on mobilemalware spread modeling. IEEE Access
10. Spedicato G (2017) Discrete time Markov chains with R. R J
11. Ferreira S, Cota W (2017) Optimized gillespie algorithms for the simulation of Markovian epidemic processes on large and heterogeneous networks. Comp Phys Commun. http://dx.doi.org/10.1016/j.cpc.2017.06.007

Fostering Adoption of Digital Payments in India for Financial Inclusion: Policies and Environment for Implementation

Aditi Bhatia-Kalluri

Abstract This paper analyses the regulations for digital payment policies that are assessed and amended based on any gaps regarding its penetration nationwide. The demonetization of the Indian banknotes led to the facilitation of various modes of digital payment. While adopting these modes, the citizens faced perils of surcharges, the inconvenience of non-real-time transactions, and stagnation of cash in the digital wallet without interest and more. Such as credit and debit cards requiring point of sale (PoS) terminals with high operational cost impeding their adoption for micro and small merchants. Hence, upon remonetization, a surge in cash usage was noticed because of the convenience of transparency it provides. The key policymakers such as the Reserve Bank of India (RBI), the Ministry of Electronics & Information Technology (MeitY), and the National Payment Corporation of India (NPCI) aim to lead India towards a less-cash society and sustain digital payment adoption. The policies and regulations around digital payments should strive to provide access to user-friendly and cost-effective financial service mobile applications to empower both merchants and consumers with a stable digital payment infrastructure. To ease the challenge for users of having to experiment and choose from various modes of digital payments, the government has mandated unified payment interface (UPI) system that consolidates the digital payment experience to promote a low-cost QR code payment acceptance solution. The paper provides policy solutions by recommending policies to ease the adoption of digital payments for the financial inclusion of every citizen and its long-term sustenance.

Keywords Financial exclusion · Low-cost payment acceptance solution · Digital policies · Information policies

A. Bhatia-Kalluri (✉)
Faculty of Information, University of Toronto, Toronto, ON 08544, Canada
e-mail: aditi.bhatia@mail.utoronto.ca

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_54

1 Introduction

Digital India Programme was launched in 2015 by the Ministry of Electronics & Information Technology (MietY) under the Government of India. As per the mission statement, the goal of the *Digital India Programme* is transformation of the digital environment to help with infrastructural development. *Digital India Programme* aims to bridge the digital divide between rural and urban by providing access to high-speed broadband, WiFi hotspots, and digital literacy, particularly, to deliver the public services digitally, that is, e-governance. In September 2016, India's leading cell phone provider, Reliance Jio launched 4G LTE networks offering practically unlimited data plans for about $6 CAD a month [15]. Although the data-phone plans were already affordable in India, this price was one of the lowest data tariffs in the world. For over a decade, the non-branded handsets have been prominent within the lower socioeconomic masses, which are significantly cheaper than the name brands yet claimed to have many of the same multimedia functions. These reliable yet lower in cost while also supporting all the features as other smartphones allow for accessible devices along with the affordable network services.

With *Digital India Programme* contributing towards the digital empowerment of the citizens, the subsequent goal of the Government of India is to foster a less-cash society for a seamless digital payment experience. This goal is following the Digital India slogan 'Faceless, Paperless, Cashless' [6]. In November 2016, the Government of India demonetized all ₹500 and ₹1000 banknotes. In a country with an estimated 98 percent of cash transactions and most of the population lacking bank accounts, the overnight discontinuation of banknotes, caused cash shortage resulting in an immediate shift to digital payment services [7]. Demonetization came across as a revolutionary economic policy measure that forced the population to adopt digital payments as an alternative to cash and mandated each citizen to own a bank account and link it to their *Aadhar* number, a universal biometric identification system for every citizen [24]. With this, *Ministry of Electronics & Information Technology* (MietY) is supervising a committed 'Digidhan Mission' (Digi = Digital; dhan = wealth), responsible to form strategies to promote and create awareness for various modes of digital payments in collaboration with all the stakeholders. These modes of payments include unified payment interface (UPI), point of sale (PoS) machines, banking cards, mobile wallets, and Internet/mobile banking (MEITY). The *National Payments Corporation of India* (NPCI), an umbrella organization for operating all retail payments in India, has also been promoting unified payments interface (UPI) mobile app that consolidates the payment methods and allows a user to instantly transfer money between bank accounts of any two parties (NPCI). Overall, a sharp jump in adoption of digital payment was noticed post-demonetization [13, 24]. According to a Wall Street Journal article, post-demonetization, millions of citizens who might not have debit or credit cards or even bank accounts were 'leapfrogging into mobile payments' instead [1]. The goal of demonetization policy and cash-less India initiative is financial inclusion for semi-urban and rural users for whom mobile phones are the only way of accessing the Internet, lifted by cheaper devices, affordable services, and faster connectivity.

2 Significance of the Problem

Various scholars such as [12, 23] propagate the fact that demonetization has necessitated digital payments, and various government regulations helped streamline the methods of payment. While enacting demonetization policy to enforce the adoption of digital payment services, the Government of India and Reserve Bank of India (RBI) also implemented several regulatory measures to ease the transition for both merchants and consumers. These included a waiver for service taxes for consumer's digital payments up to a certain amount, providing free point of sale (PoS) machines to merchants in villages, discounts on public sector services such as highway tolls, railway tickets accepting digital payments [25], and launching a UPI app [21]. The merchant discount rate (MDR) charge is a price paid by merchants to banks for accepting card payments below a certain value, the Government of India rationalized MDR applicable on debit card transactions based on the category of merchants [16] and ensured that MDR charges associated with digital payment shall not be passed to consumers [21]. While these measures encouraged the movement of the population from cash to digital modes of payments, it is critical to note that these measures were offered short-term, mostly until the discontinued currency banknotes were replaced. It is crucial to scrutinize modes of digital payments provided by the Government of India as an alternative to cash and their overall effectiveness for long-term sustenance of digital payments to foster a less-cash society.

Subsequent paragraphs, however, are indented. Post-demonetization, despite attempting to provide schemes for its proliferation, the growth for debit and credit cards have not been matched with the growing requirement for point of sale (PoS) terminals. According to RBI deputy governor, R. Gandhi, the high operating expenses of PoS infrastructure is a roadblock for its expansion [26]. Credit and debit cards are having an uneven success in India due in part to the limited number of point of sale (PoS) devices to utilize the cards. The digital transactions continued to increase after demonetization until the resurgence of the new banknotes into the economy. The remonetization of the discontinued banknotes completed in April 2017, within 5-months of demonetization in November 2016. Interestingly, near-completion of the remonetization process, the merchant's unwillingness to pay MDR charges to banks has resulted in a drop in demand for a new point of sale (PoS) devices. The remonetization led to a decline in card payment methods due to surcharge associated with their use. The merchant charges for PoS (point of sale) transactions discouraged smaller merchants from accepting electronic payments [20]. PoS payment infrastructure is particularly lacking in the rural, semi-urban areas and also for small merchants because their cost is greater than the transaction fees the market can support [3]. MDR charge and lack of widespread PoS terminals impede the ubiquity and sustenance of credit and debit card as a mode of digital payment. Thus, the credit and debit card model are failing at penetrating nationwide, and also, a resurgence of cash into the economy is making the population opt for cash rather than paying its operational costs.

While the popularity of digital wallets has remained stable, there are some impediments to its growth such as loss of interest on the money that is stagnant in a wallet, as compared to money in a bank account. There are high-transaction charges for transferring money back from a wallet to a bank account, which restricts the cash in the wallet [13]. A particular wallet brand only allows money transfer within the same wallet, which deems as restrictions when dealing with merchant tied up with another wallet [26]. This monopolization of mobile wallets by giant companies is stripping the landscape of equity in providing equal access to the users tied with other modes of digital payment. With various digital payment options, what users require is a platform that honours and embraces various modes of digital payment into one platform with maximum autonomy to transfer the funds to a wider audience through simplified means. Thus, a policy intervention for digital payments is required for financial inclusion of citizens that are rich and poor and from both urban and rural nationwide.

3 Stakeholders

The recent policy reforms such as Digital India, cashless India, demonetization of the Indian banknotes, and its remonetization in due course had major regulatory impacts. The roadblocks within digital payment policies impacting the customers and merchants to foster an inclusive digital financial environment have been highlighted. It is also pertinent to shed light onto the stakeholders and the key players: whose actions, services, or policymaking can impact the future expansion of the digital payments.

3.1 Citizens and Laymen

Customers. Impacted by the demonetization restricting cash flow in hand and looking for policies that make the adoption of the mode of digital payments easy and as transparent as cash transactions. Seeking digital mode of payments that are accepted nationwide/universally also avoids service charges and somewhat has influence on policy. Customers can indirectly influence policy through their adoption of one digital payment method over the other, which urges policymakers to help them easily adopt the mode that is more popular, while also attempting to remove roadblocks from a less accepted method, if possible.

Merchants. Initially, business impacted by demonetization due to lack of a variety of digital payment options wants to retain their customers and provides them with widely accepted payment options. Seeking digital modes of payment that avoids surcharges somewhat has influence on policy and has limited options for providing various modes of digital payment to a customer based on their company size but can choose a universally accepted digital payment mode and contribute to increasing its popularity.

3.2 Ministries, Regulatory Bodies, and Policymakers

Ministry of Electronics and Information Technology (MeitY). Incharge of the Digital India programme. Want to empower citizens with digital infrastructural development, also attempting to promote cashless India initiative. Significant influence on policy. Originators of one of the key pillars for this digital and economic policy revolution. Entrusted with the responsibility of leading this initiative on the promotion of digital transactions to create an ecosystem to enable digital payments across the country.

Department of Financial Services (Ministry of Finance). Financial inclusion is one of their key agendas. Their goals have gained exponential push with the Digital India and cashless India initiatives. Significant influence on policy. On the committee for policymaking.

Reserve Bank of India (RBI). India's Central Bank. Responsible for setting up and guiding National Payment Corporation of India for all retail payments. Payments within India are governed by the Payments and Settlement System Act, under the regulatory purview of the RBI. Significant influence on policy. Responsible for supervising, formulating, and implementing all monetary policies in India.

National Payments Corporation of India (NPCI). An umbrella organization of all retail payments in India. NCPI supervises various modes of digital payment such as RuPay, UPI, AePS, QR code payment acceptance solution, and more. Responsible for regulating various modes of digital payments. Also, formulating and amending their policies.

Cashless India. Affiliated with the digital India program. Cashless India aims to promote cybersecurity, various digital payment methods among citizens and public sectors. Somewhat has influence on policy. The initiative aims to create awareness and contribute to capacity building for a less-cash society. Significant advisors for policy formulation.

(National Institution for Transforming India) Aayog. NITI Aayog will also implement an action plan on advocacy, awareness, and handholding efforts for digital payments in the nation. Somewhat has influence on policy. The initiative aims to create awareness and contribute to capacity building for a less-cash society. Significant advisors for policy formulation.

4 Policy Solutions

While demonetization forced the population towards modes of digital payment, the remonetization showed a trend towards the use of cash again as a preferred mode of payment. This risks the redundancy of the key goals of the major economic policy altercations, which was to sustain the population's adoption of digital payments. Despite the various modes of digital payments, the perils of surcharges, monopolization of platforms persist that urge the citizens back to cash transactions. This is crucial in understanding that digital payment adoption is a long-term goal, and policy solutions for its sustenance are significant to plan.

According to a policy report by the Reserve Bank of India (RBI) released in March 2017, nine out of eleven digital payment modes set forth as cash alternatives show a decline [26]. These modes include debit/credit cards, PoS, mobile wallets, bank prepaid card, online banking, and others. With this revelation, it is important to shed light on the two modes of digital payment that remained consistently popular despite remonetization: a unified payments interface (UPI) and a quick-response code payment acceptance solution [3]. According to the RBI policy report (2017), UPI provides ease for person-to-person as well as person-to merchant transactions. It is important to note that UPI democratizes the financial landscape, rather than restricting users with a particular banking company or wallet application. UPI fosters a bank-to-bank transfer system [22]. Moreover, NPCI introduced a UPI interface owned by the government, which serves as a common app for any bank account linked at the back end allowing wider acceptance of various payment methods. UPI system provides a uniform option for anyone in India with a bank account and smartphone [3]. UPI also enables linking Aadhar number, a universal identification for every Indian citizen, which can be used for money transfer. This system is known as the *Aadhaar Enabled Payments System* (AEPS) that allows the Aadhaar number to be used for direct cash transfers, an important part of financial inclusion. UPI converges Aadhar number as a mode of payment allowing a user to pay with an Aadhaar number that offers an additional synergy with new Aadhaar-enabled bank accounts [3]. UPI supersedes the perils of previously discussed modes of digital payment by providing greater transfer limits per transaction, and the ability to transfer directly from the bank account avoids loss of interest on money that is stagnant in a wallet.

The low-cost QR payment acceptance method comes as a solution for challenge for widespread digital payment acceptance. Low-cost QR payment acceptance method is the world's first interoperable quick-response (QR) code acceptance solution developed by NPCI in collaboration with MasterCard, Visa, and American Express to expedite India's transition to a less-cash society [18]. It is a low-cost payment acceptance solution where customers pay the participating merchants by scanning a unique QR code for that business with their smartphone camera, with no other technology required on the merchant's end [3]. The scanning functionality digitizes both payments giving and accepting, which skips the processing of transactions through conventional PoS terminals. The merchant only needs to display their unique QR code at their storefront or through the acquiring bank's mobile application [18]. QR aims to standardize QR payment acceptance model nationwide. It provides customers with an opportunity to pay directly through the UPI set-up on their smartphones connected with their debit, credit, and prepaid cards [18].

The interoperability of UPI and low-cost QR payment acceptance method trumps them from the other form of digital payments [22] due to their operational cost-effectiveness, real-time money transfers, convenience, and accessibility. Their interoperability makes them prosperous. UPI app is a unified platform for customers to link their banking cards irrespective of their type, while merchant only needs to register their business account with the bank to receive payments [3]. This addresses the pain points of both the customers and the merchants by providing a unified platform that caters various bank cards with a low-cost acceptance solution. This combination specifically serves to make Digital India and the less-cash India initiatives equitable for the citizens. These two mobile payment modes have also displayed a consistent rise in value (in Rupees) and volume (number) of transactions on remonetization, while other forms have shown a decline [26]. It is important to note that the economic reform policy starting with demonetization has turned India's economy from cash-based to mobile-based within a matter of months leapfrogging the plastic money and PoS. The low-cost QR payment acceptance solution is an exemplary model for other nations looking to foster a digital payment environment. It demonstrates how a government-sponsored unified platform that is interoperable with a QR code can provide a completely digitized payment model to move beyond credit and debit cards [3]. The recent policy initiatives by the government include two promotional schemes for further adoption of digital payments: 'UPI-Referral Bonus scheme for individuals' and 'UPI-Cashback scheme for merchants,' which was valid for a year after these interoperable systems were introduced, ending on 31 March 2018. While the government recognizes these forms of mobile payments trumping the others, further recommendations on these policy solutions would help foster sustenance of a less-cash society in India.

5 Policy Recommendations

While UPI and low-cost QR payment acceptance method are payment solutions guiding a way towards complete digitization for any kind of money transfer, some pertinent policy recommendations would work as a push to feasibly foster a less-cash society in India. These recommendations highlight the benefits of digital payment awareness, incentivizing, strengthening digital infrastructure, and helping both customers and merchants choose relevant and interoperable mobile payment modes. Reforms in the regulatory regime for mobile payments should help accelerate policy actions to stabilize and sustain UPI and low-cost QR payment acceptance method to achieve scale.

- Subsequent paragraphs, however, are indented. While the UPI and low-cost QR payment acceptance method look promising, the only constraining factor for their adoption is the still incomplete adoption of smartphones, since both systems require cameras and up-to-date operating systems [3]. This calls for policy interventions by the Digital India programme to further foster the accessibility for every citizen with required infrastructural upgrades, digital literacy, affordable network services, and smartphones.
  - 'Subsidized Smartphone Scheme for low-income Citizens' in a measure to boost scalability of these modes of mobile payment.
- Micro and small merchants nationwide continue to prefer cash transactions ([4], p. 12). Adoption of a low-cost digital payment acceptance mode will give these merchants a competitive advantage to provide payment flexibility to their customers. For adoption of low-cost QR payment acceptance method by these merchants, it is significant to address their barrier to entry by raising awareness and incentivization. This calls for policy initiatives to incentivize micro and small merchants, with less than a certain amount of annual business turn-around.
  - 'New QR Merchant Cashback Scheme:' Cashback incentive scheme for micro and small merchants with newly adopted QR code. Merchants get cashback on the first certain number of payment acceptance with a minimum transaction value of a certain amount.
  - 'Promotional Scheme for a customer of a New QR Merchant:' The customers of a merchant with new QR code to get incentivized for the first transaction as a unique customer. This would encourage regular customers to motivate the vendor to adopt QR as a mode of payment acceptance.
- The application stores on mobile provide an enormous pool of applications to choose from such as various UPI apps, digital wallet apps, QR code payment acceptance apps that clutter into the application market. Users encounter difficulty to choose one app over another [10]. The regulations and policies to develop financial service applications for app stores are governed by the RBI under the *Payment and Settlement Systems Act of 2007*.
  - RBI to extend the barrier for standard requirements for launching a financial service application, which ensures its credibility and security in the marketplace. This makes the products available in the marketplace more comprehensible for users with less awareness and digital literacy allowing only plausible applications to be available as possible digital payment solutions.
  - Digital India and NITI Aayog to play a role in implicitly promoting the prominent financial service applications with proven success rate and credibility, such as UPI and low-cost QR payment acceptance method.

References

1. Abrams C, Nayak D (n.d.) Could India's cash blitz kill off cards, ATMs? The Wall Street Journal. Retrieved from https://www.wsj.com/articles/indias-cash-crackdown-prompts-more-to-pay-by-phone-1493467234 (2017)
2. Cashless India (n.d.) Retrieved from http://cashlessindia.gov.in/index.html (2019)
3. Creehan S (2017) Demonetization is catalyzing digital payments growth in India. Retrieved from https://www.frbsf.org/banking/asia-program/pacific-exchange-blog/demonetization-is-catalyzing-digital-payments-growth-in-india/ (2019)
4. Deloitte (2017) Digital payments revolution in India (Rep.). Retrieved https://bankingfrontiers.com/wp-content/uploads/2017/06/PayNext-2017-Compendium.pdf
5. DigiDhan (n.d.) Retrieved from https://digipay.gov.in/dashboard/Default.aspx (2019)
6. Digital India (2017) Digital India, ministry of electronics & information technology, government of India. Retrieved from http://digitalindia.gov.in/
7. Faden, M (2017) India's demonetization spurs digital payment services. Retrieved from https://www.americanexpress.com/us/foreign-exchange/articles/india-demonetization-digital-payment-services-growth/
8. Government of India, ministry of finance, department of economic affairs. Promotion of payments through cards and digital means. Retrieved from https://dea.gov.in/sites/default/files/Promo_PaymentsMeans_Card_Digital_0.pdf (2016)
9. GSMA, The Mobile Economy India 2016 (n.d.) Retrieved from https://www.gsma.com/mobileeconomy/india/ (2019)
10. Halan M (2017) BHIM and Bharat QR are here. Are you digital yet? Retrieved from https://www.livemint.com/Money/sY61ydhpdWuY7TT41x2bmK/BHIM-and-BharatQR-are-here-Are-you-digital-yet.html
11. Jeffrey R, Doron A (2013) The great Indian phonebook: how the mass mobile changes business, politics and daily life. Hurst, London
12. Kumar N, Puttanna K (2018) Payments transition in India–consumer preferences and policy shifts. Banks Bank Syst 13(4):17–30. https://doi.org/10.21511/bbs.13(4).2018.02
13. Malik M (2019) 11 UPI (Unified Payments Interface) Benefits—BHIM, Paytm, Google Pay, PhonePe. Retrieved from https://www.youtube.com/watch?v=-cw523kk1Hw&t=12s
14. Meity (n.d.) Modes of digital payment. Retrieved from https://meity.gov.in/modes-digital-payment (2019)
15. NDTV (2017) Cheapest Jio prepaid plans with 1GB data per day, Unlimited Calling. NDTV Business. Retrieved from www.ndtv.com/business/cheapest-jio-prepaid-plans-with-1gb-data-per-day-unlimited-calling-1783415
16. NIT (2018). Digital Payments (Government of India, NITI Aayog). Retrieved from http://niti.gov.in/writereaddata/files/document_publication/DigitalPaymentBook.pdf
17. NITI (n.d.) Government of India, National Institution for Transforming India NITI Aayog. Implementation of digital payment systems in India [Press release]. Retrieved from http://niti.gov.in/writereaddata/files/press_releases/PressRelease-MinisterCommittee.pdf (2019)
18. NPCI (2017) Government of India, National Payments Corporation of India. NPCI, Mastercard, Visa develop BharatQR [Press release]. Retrieved from https://www.npci.org.in/sites/default/files/NPCI-Mastercard-Visa-develop-BharatQR.pdf
19. Promotion of Payments through cards and digital means (n.d.) Retrieved from http://vikaspedia.in/e-governance/digital-payment/policies-and-schemes/promotion-of-payments-through-cards-and-digital-means (2019)
20. Raghavan TS (2017) Cash is back as digital payments dip on cost. Retrieved from https://www.thehindu.com/business/Industry/cash-is-back-as-digital-payments-dip-on-cost/article18458867.ece
21. RBI (2017) Government of India, Reserve Bank of India. Macroeconomic impact of demonetisation: a preliminary assessment. Retrieved from https://rbidocs.rbi.org.in/rdocs/Publications/PDFs/MID10031760E85BDAFEFD497193995BB1B6DBE602.PDF

22. Saleem SZ (2018) 4 reasons why UPI may overtake mobile wallets soon. Retrieved from https://www.livemint.com/Money/A1bTvyBsfMmZeNu6oSfozJ/4-reasons-why-UPI-may-overtake-mobile-wallets-soon.html
23. Sheetal JU, Purohit DN, Anup V (2019) Increase in number of online services and payments through mobile applications post demonetization. Adv Manage 12(1):34–38. Retrieved from http://myaccess.library.utoronto.ca/login?url, https://search-proquest-com.myaccess.library.utoronto.ca/docview/2187374735?accountid=14771
24. UIDAI (2018) Aadhaar enabled payment system | AEPS. Retrieved from https://aadhar-uidai.in/aadhar-enabled-payment-system/
25. Vadivelalagan S, Demonetization roulette: India's unusual approach to creating a cashless economy. Retrieved from https://pv.glenbrook.com/demonetization-roulette-indias-unusual-approach-to-creating-a-cashless-economy/ (March 28)
26. Waghmare A (2017) Threat to cashless economy? After demonetisation push, digital transactions recede. Retrieved from https://www.hindustantimes.com/india-news/digital-india-threatened-after-demonetisation-push-digital-transactions-recede/story-CpMaY0kcYoGLVreLhIYVHN.html
27. Yadnya (2016) UPI vs Mobile Wallets | can payment wallets survive UPI revolution? Retrieved from https://www.youtube.com/watch?v=JdyvfC01fxA&t=311s

Deep Learning-Based Adaptable Learning Analytics Platform for Non-verbal Virtual Experiment/ Practice Learning Contents

Kwang Sik Chung

Abstract Deep learning model-based learning analytics model suitable for educational and research has specific requirements. The learning analytics model is defined according to the educational requirements of the online organization and the learning operation environment of the education institutes that provide the learning analytics data. In particular, the learning analytics model is determined by the learning analytics data (learning environment operation data excluding personal information, learning content-related learning activity data, academic affair data, academic achievement data, etc.). The deep learning model-based learning analytics model of this research is developed in the form of a long-term learning analytics model and a short-term learning analytics model. Through the automatic hyperparameters tuning module of the learning analytics data management system, the long-term learning analytics model and the short-term learning analytics model produce the learning analytics results for educational institutes and individual learners. The structure and definition of the learning analytics input data and the form of output results for the long-term learning analytics model and the short-term learning analytics model are defined.

Keywords E-learning · Deep learning · Adaptable learning analytics · Virtual learning contents · Short-term learning analytics · Long-term learning analytics

1 Introduction

As the number of learners using e-learning environment (distance education) is diversified due to Covid-19 and the functions of digital learning contents and learning management systems are advanced, the personalized learning environment that should be provided to learners through the results of analysis of learners' learning activities and learning needs can be decided. Accordingly, the need for a deep learning model-based learning analytics platform for non-verbal virtual experiments

K. S. Chung (✉)
Department of Computer Science, Korea National Open University, Seoul, Korea
e-mail: kchung0825@knou.ac.kr

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_55

and practice learning contents that can be shared between countries (universities) is increasing. The definitions and concepts of learning analytics data in the e-learning environment must be shared and must support interoperability between Learning Management System (LMS).

The e-learning environment is suitable for collecting and tracking learners' learning activities data and learner-related information. Therefore, if the learning analytics platform is built on the e-learning environment, the learning analytics platform can have many advantages. In addition, if a deep learning model or a machine learning model is applied to learning analytics, high accuracy of learner's learning prediction or group learning trend prediction results can be obtained. However, a deep learning model or a machine learning model is a learning analytics data management system that can protect learner's personal information, in order to collect and track a large amount of learning analytics data for learning the deep learning model, and a distributed processing system for the deep learning model that can provide large-scale computing resources. In particular, learning analytics data for the deep learning model requires learner's participation-oriented learning contents that can generate various learning activities, and for the diversity of learning analytics data, international research exchange. It is essential to extract/analyze/manage various learning analytics data through language virtual experiments and practice learning contents.

In this chapter, we design a learning analytics model based on deep learning and classify the learning situations, learning intention, and learning goals of learners. Through a learning analytics model based on deep learning with the learning situations, learning intention, and learning goals of learners, the learning analytics aspects of learners are classified into short-term learning analytics and long-term learning analytics, and the learning analytics system for the classified short-term learning analytics and long-term learning analytics is designed. Along with the learning analytics system for the classified short-term learning analytics and long-term learning analytics, the specification and extraction module for the required learning analytics data are proposed, and the learning cloud-based virtual experiments and practice learning content system from which the learning analytics data will be extracted is also limited. And, the learning analytics system manages the accumulated learning analytics data through the learning analytics data storage, which collects, refines, and distributes the learning analytics data accumulated in the learning management system (LMS) and virtual learning contents cloud for experiments and practices where learning activities are performed.

The remainder of the paper is structured as follows: Section 2 reviews related earning analytics data management and platform, and previous learning cloud. Section 3 presents the process and design of Deep Learning-Based Adaptable Learning Analytics Platform for Non-Verbal Virtual Experiment/Practice Learning Contents. Section 4 concludes the paper by explaining the contribution of the proposed Deep Learning-Based Adaptable Learning Analytics Platform for Non-Verbal Virtual Experiment/Practice Learning Contents and mentions the limitations of the study and future research directions.

2 Related Works

In [1, 2], emotion state of a learner is used for learning analytics data and learners' emotional analytics system was proposed. In [3], learner's emotional state, learning behavior, and learning progress were analyzed for learning analytics. In [4], concept and definition of smart learning was proposed. In [5], learning cloud with combination of private cloud and public cloud was proposed. The private cloud would be constructed for private universities' learner's learning analytics data, and for public learning analytics data of universities, public cloud would be proposed. In [6], intelligent tutoring system collects and accumulates learner's learning activity and extracts their preference of learning and studying. The Learning Context Analysis System for Digital Textbook Service on Learning Cloud. In [7, 8], the primitive learning activity data is proposed. The primitive learning activity data is defined as swipe, typing, click, painting, download, save, images drawing, bookmarks, etc. Learning activity analysis for the primitive learning activity data and estimation model were combined and analyzed through data mining.

In [9–11], the concepts, requirements, and definitions of smart learning, private cloud for the personal information of distance learning universities, intelligent tutoring system for student learning activity tracing and analyses of preferences of learning and studying, and learning contents adaptation model for personalized learning contents are proposed respectively. In [12], learning activity and learner's big data were defined and collection and management methodology of learning activity and learner's big data were proposed.

3 Deep Learning-Based Adaptable Learning Analytics Platform

The learning analytics data available in LMS and the learning cloud (learning environment operation data excluding personal information, learning content-related learning activity data, academic-related environment data, academic achievement data, etc.) becomes the determining factor of the learning analytics model. The deep learning-based learning analysis model of this study is developed by division into a long-term learning analytics model and a short-term learning analytics model. The short-term learning analytics data processing system and long-term learning analytics data processing system are organically linked to extract learning analysis results for individual learners in Fig. 1. LMS, virtual learning content management server and virtual learning contents cloud for experiments and practices collect and share learning analytics data based on Experience Application Protocol Interface (xAPI).

The learner who logs into the smart learning portal server accesses the virtual learning content management server and virtual learning contents cloud for experiments and practices through learning management system (LMS). The learner

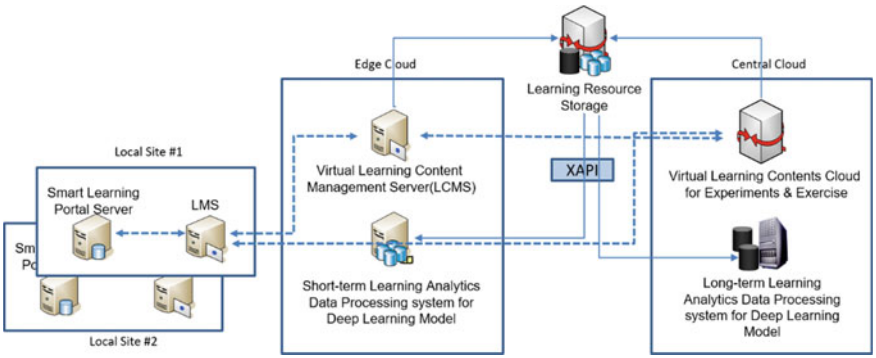


Fig. 1 Deep learning-based adaptable learning analytics platform.

performs various experiments and practice-related learning activities in the virtual learning contents cloud for experiments and practices, and these learning activities are delivered to the learning analytics data storage. This learning activity data is analyzed by the short-term learning analytics data processing system according to the learner's short-term learning goals and learning situation. The short-term learning analysis model applies a reinforcement learning model based on the Markov Decision Process (MDP) to analyze the learning analytics data of a short unit time such as a semester or an evaluation test (mid-term exam, final exam, assignment submission, etc.).

The short-term learning analysis module is built in an edge cluster environment to extract the real-time learning prescriptions for learners and real-time analysis of learning analytics data in Fig. 2. On the other hand, the long-term learning analysis model is built in a central cloud environment because a large amount of learning analytics data must be processed for the development of the learning analysis model and the learner's long-term learning prescription.

① Learners log in to the smart learning portal server (web server) using a desktop computer, smartphone, or tablet computer. Then, log on to the LMS through Single Sign-On system (SSO). And the LMS builds the learner's learning environment based on the learner's academic history and other various learning information. The LMS saves the learner's learning activities according to a set cycle.
② The LMS delivers the learner's personalized learning environment requirements to the virtual learning contents management server and learning contents cloud for virtual experiment and practice. Learners conduct virtual experiments and practices in the learning contents cloud. At this time, the learner engages in various learning activities, and continues learning. In addition, the virtual learning contents management server and learning contents cloud store the learning activities of learners according to a set period.
③ The learner's learning activity in the LMS, the learner's learning activity in the virtual learning contents management server, and the learner's learning activity

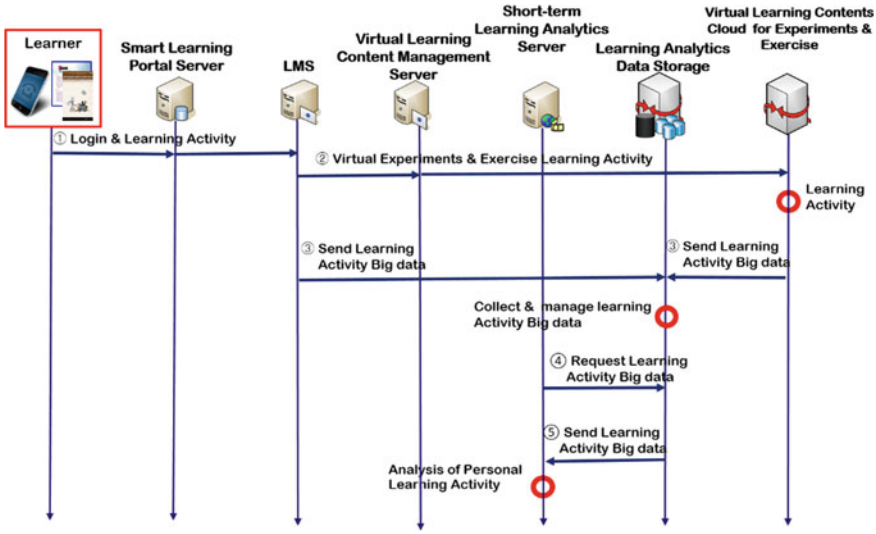


Fig. 2 The short-term learning analysis model.

in the learning contents cloud are delivered to the learning analytic storage server. The learning analytic storage server purifies, classifies, and stores learning analytics data according to the learning requirements and learning goals of the learners.

④ The short-term learning analytics server requests the learning analytics data tailored to the learner's long-term learning goals and long-term learning goals from the learning analytics storage server.

⑤ The learning analytic storage server delivers the learning analytics data suitable for the request of the short-term learning analytics server to the short-term learning analytics server. The short-term learning analytics server analyzes the learning activity of the learner according to the learning goal of the course, the teacher's learning intention, and the learner's learning ability, and saves the learning analytics result.

For then deep learning distributed processing, a container virtualization layer for using isolated resources of the CPU/GPU, distributed parallel processing platform layer to support parallel processing of personalized learning analytics models of multiple containers and deep learning model for learning analytics, and a distributed processing system stack according to the combination of the deep learning framework layers for learning analytics are built in the cloud environment. The deep learning analytics platform is built based on the deep learning distributed processing system.

The long-term learning analytics model of the deep learning-based learning analysis platform uses an ensemble-based Deep Q-Networks (DQN) model to analyze long unit time learning analytics data such as grade level or degree course in Fig. 3.

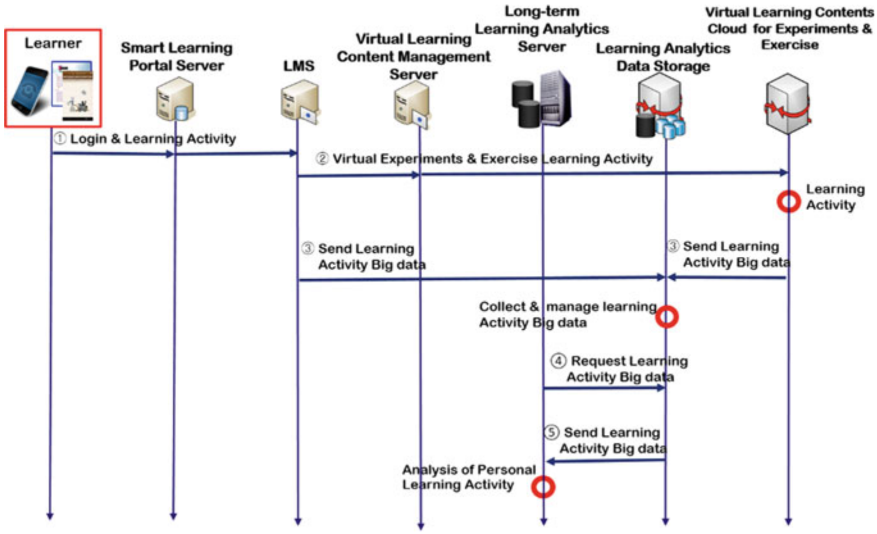


Fig. 3 The long-term learning analysis model.

① Learners log in to the smart learning portal server (web server) using a desktop computer, smartphone, or tablet computer. Then, they log in to the LMS through single-sign on system (SSO). And the LMS builds the learner's learning environment based on the learner's academic history and other various learning information, and the LMS saves the learner's learning activities according to a set cycle.

② The LMS delivers the learner's personalized learning environment requirements to the virtual learning contents management server and learning contents cloud for virtual experiment and practice. Learners perform virtual experiments and practices in the learning contents cloud. At this time, the learner engages in various learning activities, and continues learning. In addition, the virtual learning contents management server and learning contents cloud store the learning activities of learners according to a set period.

③ The learner's learning activity in the LMS, the learner's learning activity in the virtual learning contents management server, and the learner's learning activity in the learning contents cloud are delivered to the learning analytic storage server. The learning analytic storage server purifies, classifies, and stores learning analytics data according to the learning requirements and learning goals of learners.

④ The long-term learning analytics server requests the learning analytics data tailored to the learner's long-term learning goals and long-term learning goals from the learning analytic storage server.

⑤ The learning analytics data storage server delivers the learning analytics data suitable for the request of the long-term learning analytics server to the long-term learning analytics server. The long-term learning analytics server analyzes the learner's learning activities according to the learner's learning requirements and learning goals, and stores the analysis results.

4 Conclusion

As interest in e-learning contents increases due to the development of e-learning environments, the requirements for experimental and practice learning contents in e-learning environments are increasing. In addition, with the development of big data analysis technology and data collection technology, research for learner analytics in e-learning environment is being actively conducted. In this situation, the learning contents of experimentation and practice are made based on the active learning activity of the learner. Therefore, experiments and practices learning contents are a good learning environment for the accumulation and collection of learning analytics data for long-term and short-term learning analytics. In addition, since the experiments and practices learning contents provided in the virtual environment are non-verbal learning contents, they can be shared between educational institutions regardless of country, making it an advantageous environment for collecting and utilizing various learning analytics data.

In this study, the learning analysis system was designed by classifying learners' learning goals into long-term learning goals and short-term learning goals. The long-term learning analytics model is performed by combining the learning analysis results for several short-term learning goals with the learner's long-term learning goals. By using these two types of learning analytics data, an adaptable learning analytics platform for non-verbal virtual experiment and practice learning contents that can increase the accuracy of learner's learning analysis was designed.

The future research plan is to develop a learning analytics data management system for sharing individual learning data that can protect learners' personal information based on a modified Ethereum block chain that does not consume gas.

Acknowledgements This work was supported by the Korea Sanhak Foundation (KSF) in 2022.

References

1. Lee PM, Tsui WH, Hsiao TC (2015) The influence of emotion on keyboard typing: an experimental study using auditory stimuli. PLoS One 10(6):e0129056
2. El-Abbasy K, Angelopoulou A, Towell T (2015) Affective computing to enhance e-learning in segregated societies. In: Schulz C, Liew D (eds) 2015 Imperial college computing student workshop (ICCSW 2015), pp 13–20
3. Petrovica S, Pudane M (2016) Simulation of affective student-tutor interaction for affective tutoring systems: design of knowledge structure. Int J Educ Learn Syst 1:99–108
4. Chung KS, Kim YS, Lee CH, Im Jung S (2015) KNOU smart learning: beyond the future KNOU learning environment. In: AAOU 2015

5. Chung KS, Huang WHD (2013) Hybrid learning cloud platform with private cloud platforms and public cloud platforms for smart learning. In: AAOU 2015The 2013 WEI international academic conference proceedings
6. Chung KS, Kim MY (2017) Smart learning contents adaptation engine for learning devices types and learner's property for smart learning. Adv Sci Lett 23:730–734
7. Rha IJ, Im CH, Cho YH (2015) Study on learning analytics model and methodology. Seoul Metrop Off Educ
8. Jo IY, Kim Y (2013) Impact of learner's time management strategies on achievement in an e-learning environment: a learning analytics approach. J Korean Assoc Educ Inf Media 19(1):83–107
9. Chung KS, Huang WHD (2013) Hybrid learning cloud platform with private cloud platforms and public cloud platforms for smart learning. In: The 2013 WEI international academic conference proceedings, pp 30–33
10. Chung KS (2015) Design of intelligent tutoring engine for u-learning service. J Adv Inf Technol 6(2):75–79
11. Massart T, Meuter C, Van Begin L (2008) On the complexity of partial order trace model checking. Inf Process Lett 106(3):120–126
12. Jung SI, Kim YS, Lee CH, Chung KS (2018) Studies suspension prevention service of distance learning university students with learning cloud based and leaner's big data. Adv Sci Lett 24(11):7925–7929

Measuring Trust in Government Amid COVID-19 Pandemic and the Russian-Ukraine War

Nahed Azab and Mohamed ElSherif

Abstract A wealth of research emphasizes the importance of citizens' trust in public institutions. Low social capital has been proven as a substantial factor in decreasing trust in government. As a means of increasing social capital, social media is used by governments to gain citizens' trust. It is not sufficient though to create accounts on these platforms, and there should be a well-set communication strategy and a systematic mechanism to measure its success. Few studies were conducted to measure government trust in social media considering various trust dimensions. Therefore, an evaluation of the extent of trust in the government on Facebook accounts was undertaken in 2018. In that study, a framework was developed to measure trust comprising six main items: Responsiveness, Accessibility, Transparency, Effectiveness, Efficiency, and Participation. The framework was tested on a sample of the Facebook accounts of three Egyptian ministries (chosen based on their direct relation to the country's economy). After going through two major incidents, the COVID-19 pandemic, and the Russian-Ukraine war, it became pivotal to reassess trust in the government using the same framework after applying a few adjustments and considering the aspect of government trustworthiness. A comparison between both studies is discussed drawing concluded insights.

Keywords E-government · ICT · Social capital · Social media · Trust · Trustworthiness · Facebook

---

N. Azab (✉)
The American University in Cairo, P. O. Box 74, New Cairo, Egypt
e-mail: nazab@aucegypt.edu

M. ElSherif
Extend The Ad Network, New Maadi, Egypt

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_56

1 Introduction

There is a consensus among researchers and politicians on the importance of citizens' trust in their government. Several negative consequences can result from low trust in the government. For example, people show less compliance—or even resistance—to public policies and regulations [32, 99, 55], which hinder the government's ability to perform its duties [74, 43] and can even compromise on the entire legitimacy of the government [44] and the entire political system [28]. The notion of trustworthy public institutions is more crucial in nations going through a transitional democratic phase of governance [78]. Trustworthiness helps in improving government performance and communications with its citizens, which leads to effective governance [71, 85].

Low trust in government could be due to several reasons such as economic change [10], party polarization [51], or post-materialist values [45]. In addition, sometimes citizens see that the achievements of the government programs are less than their high predictions [23, 74, 79] or unacknowledged resource allocation [4]. Furthermore, insufficient social capital was found to be an integral cornerstone in reducing trust and confidence in government [49, 72].

This was the trigger that motivated us to investigate means to measure trust in government through the Facebook pages of three Egyptian ministries that have a direct effect on the economy [3], as high economic performance increases support and trust in the public institutions [23, 116]. It is though recommended to conduct a longitudinal s to assess the change in the level of trust in government as culture and values are in continuous change driving governments to be flexible and responsive to the public [99]. Longitudinal evaluation of trust is even more required especially after going through unique occurrences such as the pandemic that started in 2019 and the Russian-Ukrainian war [46]. Such incidents have evidently affected the economy in countries worldwide.

COVID-19 had a negative effect on international trading and led to a disruption in the global supply chain in different industries [111], which increased the prices of consumer goods [17]. Moreover, the Russian-Ukraine war had affected the world economy and not just Russia's resulting in high inflation, low investment, uncertainty, supply chain inefficiencies, etc. [62]. Trust is even more compromised here because it depends on how the government is dealing with these events and communicating with citizens. What makes it more challenging is the increasing amount of misinformation on social media which adds more responsibility to the government to increase responsiveness and transparency [2, 46]. During these crucial situations, gaining citizens' trust is even more vital so that the public could comply with government rules and appeals such as taking vaccinations, wearing masks, reducing panic buying [9, 54, 91].

Dealing with this critical situation requires a well-crafted communication strategy from the part of the government to acquire the trust of the public [69]. A continuous dialogue with citizens is necessary to reduce information asymmetry, encourage people to follow government policies, and show the measures undertaken by the government that meet citizens' expectations [57, 58]. Therefore, the same fundamentals we identified to measure trust (Responsiveness, Accessibility, Transparency, Effectiveness, Efficiency, and Participation) are becoming more essential.

A body of research revealed that there is a deterioration of citizens' trust in their governments [47, 22]. For example, OECD [76] revealed that the percentage of citizens in OECD countries that trust their government does not exceed 43%. However, during challenging times such as wars, pandemics, and natural disasters, the level of trust in institutions is usually affected [95, 109]. On one hand, it can increase as people when feeling external threats. They trust more their government since they do not have alternatives or sometimes people tend to rely more on strengthening in group relations when facing a common disaster and on tending to collaborate to reach better output [84, 105, 35, 34]. On the other hand, such critical events can lead to suspicions and to the fueling of conspiracy theories [110, 117]. Remarkably, some studies confirmed that citizens show a prominent level of trust in the initial stages of a disaster, but this level decreases over time [6, 84]. A possible interpretation of this change could be due to the underlying negative economic consequences such as inflation, unemployment [64, 97]. It is also worth noting that trust in government varies across countries due to cultural, social, economic, and governance differences [29, 77].

It is therefore recommended to conduct in-depth studies in each country to measure trust in public institutions and to monitor its level over time. Investigating trust is further required in case of remarkable incidents that took place over the recent period such as the COVID-19 pandemic and the Russian-Ukraine war. Hence, this paper starts by presenting the literature relevant to social capital, trust and trustworthiness of government, and how social media could strengthen them. Next, the paper presents our research methodology and the adjustments that we made in the trust metrics of the framework that was developed in our previous study in 2018. This unified framework measures social media by applying data analytics on a sample of Egyptian government Facebook accounts. A comparison between both studies is presented along with the findings discussion, conclusion, shortcomings, and suggested further research directions.

2 Literature Review

2.1 Trust and Trustworthiness

Trust and trustworthiness are usually mixed and understood as reflecting the same concept [39, 100]. Although strongly interrelated, trust and trustworthiness have two different meanings. While trustworthiness determines the characteristics or features of a trustee, trust is concerned with the perception of others of the motivation and ability of the trustee to perform a task [48, 19, 118, 37, 47]. Therefore, trustworthiness is considered as a component of trust between two different parties [24, 41, 87].

Trust between individual and social institutions is defined as institutional trust [41, 78]. It concerns the rules, roles, and norms of the institution rather than people in charge working in this institution or past communication Zucker's [119], driving them to appropriately accomplish their duties. O'Hara [75] sees that institutional trust is characterized by objectivity and by the power and authority of the institution to exert sanctions. Smith [99] argues that this understanding of institutional trust does not differentiate between institutional trust and trustworthiness and that their possible changing inter-relations over time. Even though, such drawbacks could be addressed through our early assumption that institutional trustworthiness is part of the trust bigger framework. Smith [98] identified different possibilities of trust: granted and misdirected, granted and correctly placed, or properly or incorrectly trust which may be correctly denied. These four possibilities show the difference between how the trustor perceives the trusted and the trustworthiness of the trusted.

There is though shortage in the literature investigating institutional trustworthiness compared to the one that focuses on institutional trust [99]. Regarding trust in government, one cannot deny the value of the elements of trustworthiness in setting theories related to trust in government [56, 40].

Skepticism about the honesty and trustworthiness of the government could compromise the support for a ruling system and could negatively affect the public culture [28, 116]. Therefore, we decided to consider some aspects of trustworthiness in our study that could match social media analytics as a starting point in measuring the trustworthiness of government institutions in addition to trust. The literature pointed to the measurements components of trustworthiness being ability, benevolence, and integrity [61], competence, benevolence, and honesty [106], or competence, benevolence, and integrity [36]. All components revolve around three notions: the proficiency of the government along with its care and values of fairness.

2.2 Relation Between Social Capital and Trust in Government

The concept of social capital started in 1890 but gained more importance in disciplines such as politics and social sciences since the second half of the 1990s [31]. Interest in social capital increased more with the widespread on the Internet [38]. It concerns the relationships between individuals and groups and the mutual benefits and trustworthiness that these relationships produce [83]. Social capital generation has a positive impact on education, reducing crime rates, boosting the economy, and improving government performance [49].

Researchers categorized social capital into three main types: bonding, bridging, and linking [93, 83]. Bonding capital—the strongest one that nurtures homogeneity represents the tie between friends, family, and others having similar values or circumstances. The bridging type, which forms a weaker connection compared to bonding capital, entails the liaison between friends of friends, and examples of such relationships can be seen between individuals that were together at school, college, work, etc. The significance of bridging social capital lies in its value in creating links between diverse groups contributing to better social inclusion [93]. The third form of social capital, linking, reflects the connection between the public and their government policymakers increasing resources in wider societies.

When it comes to the association between social capital and trust, one can find different notions in the literature. Several researches measure bonding and bridging social capital through trust and other parameters like citizenship customs and membership of an organization [94, 30]. Trust in public institutions can be seen as an indication of linking social capital (e.g., [65, 30]), can be impacted by social capital in general (e.g., [49, 94]), or influences trust between people and social norms [52]. We can then conclude that there exists a solid association between trust and social capital and that participation and effective governance could be reached through social capital reinforcement [43]. During the pandemic, the increase in social capital contributes to empowering the government and ensuring more endorsement from the public regarding the state's policies, measures, and actions [57]. Therefore, exploring social capital is crucial in assessing the trust and trustworthiness of the public sector. We claim in the context of this research that the linking type of social capital can be produced over the social media pages of the government.

2.3 Can ICT Increase Trust in Government?

Governments perceive e-government as a powerful application mechanism that can reestablish trust because it enables the government to publish information about the performed achievements. In addition, e-government creates a smooth interaction between citizens and the public sector [86] portraying an image of a high-responsive government [104]. Moreover, e-government can improve the service delivered to citizens, improve the internal efficiency of public administration, and encourage civic participation [81, 47]. E-government is primarily seen as a mean to develop trustworthy institutions and to build trust in government [115, 104, 90]. Citizens tend to interact with government entities and take part in policy formulation [18]. This kind of citizen commitment is a major element of social capital, which proved to be pivotal in reinforcing trust in government [49]. High social capital is even more required in critical situations to ensure public compliance with the measures and actions set by governmental institutions [57].

Several research concluded that citizens' experiences with e-services are strong indicators of trustworthiness cues that would reinforce the citizens' trust in the public sector [98]. However, websites that provide these services cannot fully fulfill citizens' requirements and gain their trust. Interestingly, the literature confirmed that trust and e-government are bidirectional, there is no doubt that efficient e-government adoption would strengthen confidence in public institutions, and in their future performance, e-services will only be used if people believe that government institutions are trustworthy [81, 114, 104, 7, 34]. Government trustworthiness is even more vital in increasing the usage of e-services that cannot sometimes compete with those provided by companies (e-commerce), and trustworthy public institutions would reduce citizens' need to depend on their traditional visits and physical communications with government entities [102]. Focusing on trustworthiness will draw the attention of government management toward further trust-building activities and can be better controlled and guided by public decision-makers [19, 118]. Furthermore, Whiteley et al. [116] noted that policy outcome (delivering a good quality service) is one pillar of trust along with the policy process (i.e., the trustworthiness of the service provider). Citizens' perceptions toward the fairness, transparency, and info-oriented policy process proved to be essential in building trust in government [13, 15]. Understanding the policy-making process and ensuring the honesty and trustworthiness of decision-makers would drive people to abide by policies that could sometimes be against their interests [59].

During critical incidents (such as the pandemic and the Russian-Ukraine war), transparency, responsiveness, and engagement in the decision process are further required to increase social capital, improve government trustworthiness, and win public trust, which would evidently enable public policymakers to better manage any crisis and its economic implications [89, 96, 47, 27, 95, 46, 57]. These features cannot be fully provided by e-services, and public websites are primarily used to disseminate information and to offer services and not do not allow sufficient room for interactivity or engagement [101, 43]. It is important to direct the government's efforts toward communicating with the public to shape their perception of a caring, responsive, and honest government. Smith [99] and Houston and Harding [43] urged further exploitation of ICT capacity to enhance transparency, direct contact, and responsiveness. Social media could hence assist the government in improving its trustworthiness due to the features they possess allowing for better openness and visibility of public employees. Due to their contribution to increasing social capital, governments are increasingly recognizing the benefit of social media in increasing citizens' trust and are adopting them as part of their communication projects [8].

Social media would also assist the government in addressing misinformation that propagates exponentially during crises. Being exposed to more misinformation during crises minimizes people's both declarative and procedural context-specific knowledge about the responsive measures and policies taken by the government leading to less trust in government [11, 112, 46]. During the pandemic, social media helped in spreading misinformation being the first and sometimes main source of information that connects people with family and friends. Daily newspapers or government web pages are not frequently visited and do not have a high influence on citizens. Being exposed to a high volume of misinformation demotivates social media users to seek correct information from official sources [112]. Thus, an effective presence of the government in the same social platforms used by citizens would largely increase the trust in government performance, values, and beliefs and combat the credibility of misinformation.

3 Methodology

Facebook continued to be the dominant social network for the Egyptian public authorities. The selected Facebook pages that directly relate to the economy of Egypt were the Ministry of Tourism Egypt (MOT), the Ministry of International Cooperation (MIIC), the Ministry of Trade and Industry (MIFT), and the Egyptian Tourism Authority (Experience). In the last paper, we carefully examined the previous literature to identify the most common trust dimensions. Then, we added our perceptions to match the ministries' Facebook accounts. We focused on Facebook because of its importance and citizen availability. In that year, 2018, we had access to Crowd Analyzer, a software tool, for accounts' performance and sentiment analysis using its Artificial Intelligence algorithms (AI). All the Facebook pages selected in our previous were also directly related to the economy of Egyptian. Moreover, at that time, we dropped the Ministry of Tourism Egypt (MinistryofTourismEgypt) not been active since January 2017. However, in this paper, we included the Ministry of Tourism and Antiquities (MOTA) because it has been active.

Since then, many privacy issues were aggressively revisited by Facebook, and many tools stopped crawling data for analysis [66]. In addition, Crowd Analyzer has been disabled in 2021. At the time of creating the paper in 2022, we could not monitor social media accounts automatically and extract their information including posts and comments as in 2018. Therefore, we have used an alternative tool, Socialbakers, which is a leader in social accounts' performance. Socialbakers help in measuring some of the items that would have been difficult to do manually. Moreover, we were able to add more information that was not possible before such as:

- The number of times users have shared any post by the ministry with their friends.
- The number of new page followers compared to the previous period.

The monitoring period was from January 2022 to the end of May 2022.

4 Proposed Framework

The six dimensions proposed in our last study [3] were: Responsiveness (ten items), Accessibility (five items), Transparency (two items), Effectiveness (three items), Efficiency (four items), and Participation (six items). While the six dimensions used in the previous study remained the same, some sub-items have changed due to the APIs' limitation forced by Facebook after the Cambridge Analytica (https://www.tandfonline.com/doi/abs/10.1080/21670811.2019.1591927) data scandal.

The sub-items that we changed are:

- Responsiveness—Exclude (Reply Rate): The only viable way to calculate the reply rate of the page is to have access to its insights with a username and password. This is no longer an option to analyze on public pages.

- Participation—Exclude (People Engaged): All tools have no more extended access to peoples' names. Therefore, it is impossible to count the number of unique people engaged.
- Participation—Exclude (Sentiment Score): Tools cannot capture public comments unless the tools have private access with a username and password.
- Participation—Exclude (Engagement Rate): Since we are not able to identify the individual unique users, we cannot calculate the Engagement Rate.
- Participation—Updated (Interactions/post): Using Socialbakers, we can capture the number of times that users have shared the content. We have added this to the number of likes and comments to get a clearer number of users' participation.
- Participation—Include (Net followers): The number of net followers gained in the time period that was monitored.
- Participation—Include (Followers growth): We have compared the number of followers from January to May 2022 with August to December 2021.

The framework maintained its solid structure with six areas. However, the sub-items changed as highlighted. Responsiveness (nine items—Res.): the readiness of the page and their attendance to the online users. Accessibility (two items—Acc.): their connection with other governmental pages. Transparency (two items—Tra.): Are they sharing their agenda and allowing users to post to the page? Effectiveness (three items—Effe): Do they solve users' problems and raised issues? Efficiency (four items—Effi): How fast and accurately do they solve the issues? Participation (five items-Par.): The number of page-likes and engagement over time with its content.

Table 1 highlights all the dimensions with the new and updated sub-items and their scoring methods:

While the binary and normalized scores remained the same, we have introduced a new scoring mechanism for the number of followers. The net number of followers is the total number of followers gained or lost in a given time. In our study, if the number of new followers is less than the number of unfollows, the total net will be negative. The followers' growth is as follows:

((Net followers in the given timeframe)-(net followers in the previous timeframe))/(net followers in the previous timeframe).

If the number of net followers in the previous period is higher than in the current period, the percentage will be negative. Of course, it is not necessary to have the highest follower growth if you have the highest number of net followers. While the number of net followers shows user trust in the page by following it, the growth percentage shows the trust over time. If it is declining or more users are following than before.

Stemming from the fact that trustworthiness has become pivotal to e-government research [16], we have explored the study of Janssen et al. [47] that investigated the trustworthiness of e-government by deriving a comprehensive theory through interpretive structural modeling. In this study, Janssen et al. [47] highlighted 20 factors affecting citizens' perceptions of e-government trustworthiness. While reading it, we discovered the following: 1—items might be difficult to measure using tools. This will need further investigation. 2—items that are different from our study, and we can obtain them by manually examining the comments or discovering advanced

Table 1 Trust dimensions and corresponding sub-items

Item	Sub-item	Scoring method
Res	Verified	Y = 1 N = 0
	Listed	Y = 1 N = 0
	Phones	Y = 1 N = 0
	Emails	Y = 1 N = 0
	Address	Y = 1 N = 0
	CTA button	Y = 1 N = 0
	Reply to comments	Y = 1 N = 0
	Reply to messages	Y = 1 N = 0
	Post to page	Y = 1 N = 0
Acc	Page Liked	Y = 1 N = 0
	Related posts	Y = 1 N = 0
Tra	Content	Y = 1 N = 0
	Approval	Y = 1 N = 0
Effe	Relevant info	Y = 1 N = 0
	Problem solved	Y = 1 N = 0
	Complete info	Y = 1 N = 0
Effi	Automated messages	Y = 1 N = 0
	Problem solved	Y = 1 N = 0
	Created apps	Y = 1 N = 0
	Fast reply	Instantly = 1, in minutes = 0.8, within an hour = 0.6, in few hours = 0.4, within a day = 0.2
Par	Citizens' input	Y = 1 N = 0
	Citizens' meet	Y = 1 N = 0
	Interactions/post	Normalized score (0–1)
	Net followers	Ranked the pages from 0 to 3 according to the number. The lowest is 0
	Followers' growth	Ranked the pages from 0 to 3 according to the number. The lowest is 0

analysis tools. 3—items that match our study. We have highlighted the dimension and sub-item matching in this case. 4—items that cannot be obtained using the current tools. Table 2 shows the 20 factors that affect e-government identified by Janssen et al. [47].

Table 2 Twenty factors that affect trust Janssen et al. [47]

Item	Notes
Trust of government	Difficult using social media analytics
Trust of Internet	Difficult using social media analytics
Disposition to trust	Difficult using social media analytics
Perceived risk	Can be obtained with manual examination/advanced tools
Privacy concerns	Can be obtained with manual examination/advanced tools
Perceived security	Can be obtained with manual examination/advanced tools
Political attitudes	Can be obtained with manual examination/advanced tools
Transparency	We have this under effectiveness - > relevant info
Perceived prior knowledge	We have this under transparency - > content (the page shares information about its events, agenda, meeting results, etc.)
Accountability	We have this under effectiveness - > problem solved
Responsiveness	We have this under efficiency - > automated messages and problem solved
Service quality	We do not have this anymore because we cannot capture users' comments for now
Satisfaction	We do not have this anymore because we cannot capture users' comments for now
System quality	We do not have this anymore because we cannot capture users' comments for now
Perceived ability to use	Difficult using social media analytics
Use	We have this under the participation
Benevolence	Can be obtained with manual examination/advanced tools
Integrity	Can be obtained with manual examination/advanced tools
Competence	We have this under responsiveness
Trustworthiness of e-government	Difficult using social media analytics

We are exploring new tools that might provide further analysis of some missing items such as sentiment score. The new tools do not depend on Facebook APIs but rather use Web Scraping techniques to capture the comments. Then, using their NLP techniques, they can provide automated sentiment analysis.

5 Findings and Discussion

During the timeframe, January 2022–May 2022, we captured 1714 posts made by the ministries which gained 325,930 reactions, 54,958 shares, and 45,503 comments. This is much higher than our previous study, which captured only 199 posts. The following table (Table 3) shows the analysis of all six dimensions for the four pages. It has both old and new sub-items for ease of comparison as well.

Comparing the last study to this one, we noticed that the MIIC switched last place with MIFT. MIFT scored a higher rate in this study than in 2018, while experience scores were higher in both studies. Overall, the ministry with the highest trust score compared to other ministries is the Ministry of Tourism and Antiquities. It is difficult to compare the progress over time with the total score because the sub-items have changed. However, after comparing the common 22 sub-items out of the total 30 sub-items, we noticed a huge increase in MIFT score from 8.6 to 13.4. MIFT page started to reply to comments, allowed timeline posts, started to post more related posts with more information, and started automated reply messages. Also, we noticed a small drop in experience Facebook page. The page was used to answer questions more efficiently in the previous study than in this study. This indicates that the pages had almost the same trust level as the previous study in all dimensions except for participation which included new sub-items.

We will discuss every dimension and its noticeable results across all ministries. First, under responsiveness, all ministries are verified and have contact addresses. MICC got the lowest score not being listed or having email and call to action button. It does not also reply to comments. On the other hand, Ministry of Tourism and Antiquities was the best scoring 9/9. Second, accessibility, all ministries, except Eg.ExperienceEgypt, liked other ministries' pages and posted about other ministries' common programs.

Third, under transparency, Eg.ExperienceEgypt is focused on Egypt's experiences with no posts regarding its agenda or events. However, all other ministries shared posts regarding their events. All posts sent to the ministries' pages have to be approved by the ministry before being shown to the public. We have found only one wall post by a user that was on a page wall which might indicate that no posts are accepted, or no users were using this feature with ministries.

For the fourth dimension, effectiveness, we have analyzed users' comments manually and how the ministries engaged with them. All ministries have sent engaging content that users found relevant to the government's mission. However, again, we find that MIIC and MIFT lag with no effective replies to users' concerns or questions. Unlike this, MOTA and experience answered effectively as shown in the following sample pictures. Their replies were to the point solving any users' concerns (Fig. 1):

The next dimension, efficiency, measured the level of service. We noticed MIIC and MIFT were the only ministries enhancing the speed of service using automated messages. Lastly, users' participation was a magnificent performance by Eg.ExperienceE- gypt which scored 9/9. The ministry had the highest number of interactions with 106,534 reactions, 13,015 shares, and 15,558 comments. During the timeframe, the page had 249,340 new followers which were 500% more than the previous period.

Table 3 Findings of old and new studies

Dimensions	Sub-items	MIIC		MIFT		Experience		MOTA	
		Old	New	Old	New	Old	New	Old	New
Res	Verified	0	1	1	1	1	1	NA	1
	Listed	1	0	1	1	1	1	NA	1
	Phones	1	1	1	1	0	0	NA	1
	Emails	1	0	0	0	1	1	NA	1
	Address	1	1	1	1	1	1	NA	1
	CTA button	0	0	1	1	1	1	NA	1
	Reply to comments	0	0	0	1	1	1	NA	1
	Reply to messages	1	1	1	1	1	0	NA	1
	Reply rate	0	NA	0	NA	1	NA	NA	NA
	Post to page	1	1	0	1	0	1	NA	1
	Total	6	5	6	8	8	7	NA	9
Acc	Page liked	1	1	1	1	0	0	NA	1
	Related posts	0	1	0	1	0	0	NA	1
	Total	1	2	1	2	0	0	NA	2
Tra	Content	1	1	1	1	1	0	NA	1
	Approval	0	0	0	0	0	0	NA	0
	Total	1	1	1	1	1	0	NA	1
Effe	Relevant info	1	1	0	1	1	1	NA	1
	Problem solved	0	0	0	0	1	1	NA	1
	Complete info	0	0	0	0	1	1	NA	1
	Total	1	1	0	1	3	3	NA	3
Effi	Automated messages	0	1	0	1	0	0	NA	0
	Problems solved	0	0	0	0	1	0	NA	0
	Created apps	1	0	0	0	0	0	NA	0
	Fast reply	0.6	0.2	0.6	0.4	1	0.2	NA	0.8
	Total	1.6	1.2	0.6	1.4	2	0.2	NA	0.8
Par	Citizens' input	0	0	0	0	0	1	NA	1
	Citizens meet	0	1	0	0	0	1	NA	1
	Likes/post	1	NA	0.2	NA	0.27	NA	NA	NA

(continued)

Table 3 (continued)

Dimensions	Sub-items	MIIC		MIFT		Experience		MOTA	
		Old	New	Old	New	Old	New	Old	New
	Comments/ post	1	NA	0.79	NA	0.54	NA	NA	NA
	People engaged	0.375	NA	1	NA	0.73	NA	NA	NA
	Sentiment score	0.6	NA	0	NA	0.85	NA	NA	NA
	Interactions/ post	NA	0.04	NA	0.03	NA	1	NA	0.42
	Net followers	NA	0	NA	1.00	NA	3	NA	2
	Followers growth	NA	0	NA	2.00	NA	3	NA	1
	Total	2.975	1.04	1.99	3.03	2.39	9	NA	5.42
Overall score		13.575	11.24	10.59	16.43	16.39	19.2	NA	21.22

Fig. 1 MOTA and experience answered

6 Conclusion

The decline in trust can negatively affect how citizens see the government and use its e-services to address uncertainty, anxiety, and risk that could be more eroded during crises [63, 5]. Risk communication strategies are consequently recommended that involve catering risk clarification to different recipients, respecting their values, and promoting communal and individual decision-making [68]. Public administrators cannot play behind the scenes anymore, and risk communication dictates also that the provision of timely, relevant, and accurate information generates trust in governmental responses [27]. Reforming the public sector and increasing trust requires strategies not only to improve competence but also to increase government trustworthiness—through securing an effective communication channel with the government and disseminating information about government strategies and activities [14] and to exploit the features provided by information technology to encourage civic participation [43]. Public administrators and policymakers need to consider the behavioral forms of social media users as they are the most audience exposed to misinformation [46]. Misinformation proved to affect the degree of trust in public institutions and the overall compliance with government measures and policies [67, 112]. Without using technologies that encourage interactions with citizens and improve transparency and responsiveness that help in perceiving a trustworthy government (like social media), e-government initiatives could not achieve their goals and would increase bureaucracy "transforming street-level bureaucracy into screen-level bureaucracy" [12].

Despite the growing interest of governments in gaining citizens' trust [82] and the wealth of literature about the concept of trust in public administration, there are still limited empirical studies on this topic [20, 47]. Trustworthiness, therefore, is gaining more interest in e-government research [16, 114], but there are still unexplored areas related to additional components of e-government trustworthiness [16, 88, 98, 99].

This paper, thus, aims to address a less investigated research field related to a systematic evaluation of trust and trustworthiness in social media government pages. It developed a measurement framework based on data collected from these platforms. The framework was tested over two time periods—in 2018 and 2022 on the Facebook pages of four Egyptian ministries whose activities impact the country's economy. Both studies examined the trust dimensions (Responsiveness, Accessibility, Transparency, Efficiency, and Participation); however, we carried out some adjustments in the second one due to the privacy change policies on Facebook and after incorporating additional items that measure trustworthiness. The comparative analysis between both studies showed that the ranking changed; the Ministry of International Cooperation became in last place instead of the Ministry of Trade and Industry which ranks third. Experience scores the second. The ministry with the highest trust score is the Ministry of Tourism and Antiquities.

As a research drawback, there is still no in-depth sentiment analysis of these government accounts. Future research can explore further tools that perform analysis in this area, especially in Arabic language content. Other research triangulation venues could be to conduct interviews with policymakers of government institutions and to obtain direct feedback from people through interviews and surveys to assess the existence and extent of the gap between both citizens and public institutions. Additionally, examining the change in weights of trust measurements associated with various levels of government would reveal valuable outputs.

References

1. Abdelghaffar H, Kamel S, Duquenoy P (2010) Studying e-government trust in developing nations: case of university and college admissions and services in Egypt. In: Proceedings of the international information management association conference, Utrecht The Netherlands
2. Alam MA (2020) Leading in the shadows: understanding administrative leadership in the context of COVID-19 pandemic management in Bangladesh. Int J Publ Leadersh 17(1):95–107
3. Azab N, ElSherif M (2018) A framework for using data analytics to measure trust in government through the social capital generated over governmental social media platforms. In: The 19th annual international conference on digital government research (dgo.2018), May 30th June 1st, 2018
4. Baldassare M (2000) California in the new millennium: the changing social and political landscape. University of California Press, Berkeley
5. Balog-Way DHP, McComas KA (2020) COVID-19: reflections on trust, tradeoffs, and preparedness. J Risk Res 1–11
6. Bangerter A, Krings F, Mouton A, Gilles I, Green E, Clemence A (2012) Longitudinal investigation of public trust in institutions relative to the 2009 H1N1 pandemic in Switzerland. PLoS ONE 7(11):e49806
7. Belanger F, Carter L (2008) Trust and risk in e-government adoption. J Strat Inf Syst 17(2):165–176
8. Bertot J, Jaeger P, Grimes J (2012) Promoting transparency and accountability through ICTs, social media, and collaborative e-government. Transforming Gov People Process Policy 6(1):78–91
9. Blair RA, Morse BS, Tsai LL (2017) Public health and public trust: survey evidence from the Ebola virus disease epidemic in Liberia. Soc Sci Med 172(1–2):89–97
10. Bok D (1997) Measuring the performance of government. In: Nye J, Zelikow P, King D (eds) Why people don't trust government. Harvard University Press, Cambridge, MA
11. Bose R (2004) Knowledge management metrics. Ind Manage Data Syst 104(6):457–468
12. Bovens M, Zouridis S (2002) From street-level to system-level bureaucracies: how information and communication technology is transforming administration discretion and constitutional control. Public Adm Rev 62(2):174–184
13. Cain B, Russell E, Dalton J, Scarrow SE (2003) Democracy transformed? Expanding political opportunities in advanced industrial democracies. Oxford University Press, Oxford
14. Campbell AL (2003) How policies make citizens: senior political activism and the American welfare state. Princeton University Press, Princeton, NJ
15. Carman C (2010) The process is the reality: perceptions of procedural fairness and participatory democracy. Polit Stud 58:731–751
16. Carter L, Belanger F (2005) The utilization of e-government services: citizen trust, innovation and acceptance factors. Inf Syst J 15(1):5–25
17. Casper H, Rexfors A, Riegel D, Robinson A, Martin E, Awwad M (2021) The impact of the computer chip supply shortage. In: Proceedings of the international conference on industrial engineering and operations management Bangalore, India, 16–18 Aug 2021

18. Chang A, Kannon K (2008) Leveraging web 2.0 in government. E-Government technology series, IBM center for the business of e-government: Washington, DC, USA. Online Jan 2018 at http://www.businessofgovernment.org/sites/default/files/chang_fall08.pdf
19. Cho YJ, Lee JW (2011) Perceived trustworthiness of supervisors, employee satisfaction and cooperation. Public Manage Rev 13(7):941–965
20. Cho YJ, Poister TH (2013) Human resource management practices and trust in public organizations. Public Manage Rev 15(6):816–838
21. Christensen T, Laegreid P (2005) Trust in government: the relative importance ofservice satisfaction, political factors, and demography. Public Perform Manage Rev 28(4):487–511
22. Citrin J, Luks S (2001) Political trust revisited: déjà-vu all over again? In: Hibbing JR, Theiss-Morse E (eds) What is it about government that Americans dislike? Cambridge University Press, New York, pp 9–27
23. Clarke HD, Sanders D, Stewart MC, Whiteley P (2004) Political choice in Britain. Oxford University Press, Oxford
24. Coleman JS (1994) Foundation of social theory. First Harvard University Press, Cambridge, MA
25. Colesca S (2009) Increasing E-Trust: a solution to minimize risk in e-government adoption. J Appl Quant Methods 4(1)
26. Crowdanalyzer.com; The 1st Arabic focused internationally recognized social media monitoring platform, Online Nov 2017 at http://crowdanalyzer.com
27. Deslatte A (2020) The erosion of trust during a global pandemic and how public administrators should counter it. Am Rev Public Adm 50(6–7):489–496
28. Easton D (1979) A systems analysis of political life. University of Chicago Press, Chicago, IL
29. Edelman (2022) 2022 Edelman trust barometer. Online Oct 2022 at https://www.edelman.com/trust/2022-trust-barometer
30. Ekici T, Koydemir S (2014) Social capital, government and democracy satisfaction, and happiness in turkey: a comparison of surveys in 1999 and 2008. Soc Ind Res 118(3):1031–1053
31. Ferragina E, Arrigoni A (2016) The rise and fall of social capital: requiem for a theory? Polit Stud Rev 15(3):355–367
32. Gamson WA (1968) Power and discontent, vol 124. Dorsey Press, Homewood, IL
33. Gauld R, Gray A, McComb S (2009) How responsive is e-government? Evidence from Australia and New Zealand. Gov Inf Q 26(1):69–74
34. Goldfinch S, Taplin R, Gauld R (2021) Trust in government increased during the covid-19 pandemic in Australia and New Zealand. Aust J Public Adm 80(1):3–11
35. Greenaway KH, Cruwys T (2019) The source model of group threat: responding to internal and external threats. Am Psychol 74(2):218–231
36. Grimmelikhuijsen S, Knies E (2017) Validating a scale for citizen trust in government organizations. Int Rev Adm Sci 83(3):583–601
37. Grimmelikhuijsen SG, Meijer AJ (2014) Effects of transparency on the perceived trustworthiness of a government organization: evidence from an online experiment. J Public Adm Res Theory 24(1):137–157
38. Gudmundsson G, Mikiewicz P (2012) The concept of social capital and its usage in educational studies. Social capital and education, foundation for the development of the education system. Online Jan 2018 at https://repozytorium.amu.edu.pl/bitstream/10593/5897/1/55-80.pdf
39. Hardin R (1993) The street-level epistemology of trust. Polit Soc 21(4):505–529
40. Hardin R (2000) The public trust. In: Pharr SJ, Putnam RD (eds) Disaffecteddemocracies. Princeton University Press, Princeton, pp 31–51
41. Harre´ R (1999) Trust and its surrogates: psychological foundations of political process. In: Warren ME (ed) Democracy and trust. Cambridge University Press, Cambridge
42. Harrison T, Guerrero S, Burke B, Cook M, Cresswell A, Helbig N, Hrdinova J, Pardo T (2012) Open government and e-government: democratic challenges from a public value perspective. Information Polity 17(2):83–97

43. Houston DJ, Harding LH (2013) Public trust in government administrators. Public Integrity 16(1):53–76
44. Inglehart R (1990) Culture shifts in advanced industrial societies. Princeton University Press, Princeton
45. Inglehart R (2000) Globalization and postmodern values. Wash Q 23(1):215–218
46. Islam S, Mahmud R, Ahmed B (2021) Trust in government during COVID-19 pandemic in Bangladesh: an analysis of social media users' perception of misinformation and knowledge about government measures. Int J Public Adm. https://doi.org/10.1080/01900692.2021.2004605
47. Janssen M, Nripendra RP, Slade EL, Dwivedi YK (2018) Trustworthiness of digital government services: deriving a comprehensive theory through interpretive structural modelling. Public Manage Rev 20(5):647–671
48. Kee HW, Knox RE (1970) Conceptual and methodological considerations in the study of trust and suspicion. J Conflict Resolut 14(3):357–366
49. Keele L (2007) Social capital and the dynamics of trust in government. Am J Polit Sci 51(2):241–254
50. Khan GF, Yoon HY, Park HY (2014) Social media communication strategies of government agencies: twitter use in Korea and the USA. Asian J Commun 24:60–78
51. King D (1997) The polarization of American parties and mistrust of government. In: Nye J, Zelikow P, King D (eds) Why people don't trust government. Harvard University Press
52. Knack S, Keefer P (1997) Does social capital have an economic payoff? A cross- country investigation. Q J Econ 112(4):1251–1288
53. Laroche M, Habibi M, Richard M, Sankaranarayanan R (2012) The effects of social media based brand communities on brand community markers, value creation practices, brand trust and brand loyalty. Comput Hum Behav 28(5):1755–1767
54. Larson HJ, Heymann DL (2010) Public health response to influenza A(H1N1) as an opportunity to build public trust. J Am Med Assoc 303(3):271–272
55. Lee Y, Schachter HL (2019) Exploring the relationship between trust in government and citizen participation. Int J Public Adm 42(5):405–416
56. Levi M, Stoker L (2000) Political trust and trustworthiness. Annu Rev Polit Sci 3:475–507
57. Liu J, Shahab Y, Hoque H (2022) Government response measures and public trust during the COVID-19 pandemic: evidence from around the world. Br J Manag 33(2):571–602
58. Luhmann N (1979) Trust and power. Wiley, Chichester
59. Luskin RC, Fishkin JS, Jowell R (2002) Considered options: deliberative polling in Britain. Br J Polit Sci 32:455–487
60. Maaty A (2014) Crisis of confidence: the tantalizing hope of building trust. Online Jan 2018 at http://egyptoil-gas.com/features/crisis-of-confidence-the-tantalizing-hope-of-rebuilding-trust/
61. Mayer RC, Davis JH, Schoorman FD (1995) An integrative model of organizational trust. Acad Manag Rev 20(3):709–734
62. Mbah R, Wasum D (2022) Russian-Ukraine 2022 War: a review of the economic impact of Russian-Ukraine crisis on the USA, UK, Canada, and Europe. Adv Soc Sci Res J 9(3)
63. McKnight DH, Choudhury V, Kacmar C (2002) Developing and validating trust measures for e-commerce: an integrative approach. Inf Syst Res 13(3):334–359
64. Meltzer MI, Cox NJ, Fukuda K (1999) The economic impact of pandemic influenza in the united states: priorities for intervention. Emerg Infect Dis 5(5):659–671
65. Mendoza-Botelho M (2013) Social capital and institutional trust: evidence from Bolivia's popular participation decentralisation reforms. J Dev Stud 49(9):1219–1237
66. Meta (2018) An update on our plans to restrict data access on facebook. Online Oct 2022 at https://about.fb.com/news/2018/04/restricting-data-access/
67. Michelle Driedger S, Maier R, Jardine C (2018) Damned if you do, and damned if you don't: communicating about uncertainty and evolving science during the H1N1 influenza pandemic. J Risk Res 24(5):574–592

68. Morgan MG, Fischhoff B, Bostrom A, Atman CJ (2002) Risk communication: a mental models approach. Cambridge University Press
69. Morgeson FV III, VanAmburg D, Mithas S (2011) Misplaced trust? exploring the structure of the e-government– citizen trust relationship. J Public Adm Res Theory 21:257–283
70. Mourtada R, Salem F (2011) Civil movements: the impact of facebook and twitter. Arab Soc Media Rep 1(2):1–30. Online Jan 2018 at http://unpan1.un.org/intradoc/groups/public/documents/dsg/unpan050860.pdf
71. Murphy K (2004) The role of trust in nurturing compliance: a study of accused tax avoiders. Law Hum Behav 28(2):187–209
72. Newton K (1997) Social capital and democracy. Am Behav Sci 40(5):575–586
73. Norris P (2001) Digital divide: civic engagement, information poverty, and the internet worldwide. Cambridge University Press, New York
74. Nye JS Jr, Zelikow PD (1997) Conclusions: reflections, conjectures, and puzzles. In: Nye J, Zelikow P, King D (eds) Why people don't trust government. Harvard University Press
75. O'Hara K (2004) Trust: from Socrates to Spin. Icon Books Ltd., Duxford, Cambridge
76. OECD (2017) Trust in government. Online Jan 2018 at http://www.oecd.org/gov/trust-in-government.htm
77. OECD (2021) Building trust to reinforce democracy: key findings from the 2021 OECD survey on drivers of trust in public institutions. Online Oct 2022 at https://www.oecd-ilibrary.org/sites/b407f99c-en/index.html?itemId=/content/publication/b407f99c-en
78. Offe C (1999) How can we trust our fellow citizens?. In Warren ME (ed) Democ
79. Orren G (1997) Fall from grace: the public's loss of faith in government. In: Nye J, Zelikow P, King D (eds) Why people don't trust government. Harvard University Press
80. Panagiotopoulos P, Bigdeli Z, Sams S (2014) Citizen–Government collaboration on social media: the case of twitter in the 2011 riots in England. Gov Inf Q 31(3):349–357
81. Parent M, Vandebeek C, Gemino A (2005) Building citizen trust through e-government. Gov Inf Q 22(4):720–736
82. Park H, Blenkinsopp J (2011) The role of transparency and trust in the relationship between corruption and citizen satisfaction. Int Rev Adm Sci 77(2):254–274
83. Putnam RD (2001) Bowling alone: the collapse and revival of American community. Simon and Schuster 19
84. Quinn SC, Parmer J, Freimuth VS, Hilyard KM, Musa D, Kim KH (2013) Exploring communication, trust in government, and vaccination intention later in the 2009 H1N1 pandemic: results of a national survey. Biosecur Bioterror 11(2):96–106
85. Raab CD (1998) Electronic confidence: trust, information and public administration. In Snellen ITM, van de Donk WBHJ (eds) Public Administration in an information age. OS Press, Amsterdam
86. Ravishankar MN (2013) Public ICT innovations: a strategic ambiguity perspective. J Inf Technol 28(4):316–332
87. Reed MI (2001) Organization, trust and control: a realist analysis. Organ Stud 22(2):201–228
88. Robinson SE, Liu X, Stoutenborough JW, Vedlitz A (2013) Explaining popular trust in the department of homeland security. J Public Adm Res Theory 23(3):713–733
89. Rothstein B, Stolle D (2008) The state and social capital: an institutional theory of generalized trust. Comp Polit 40:441–459
90. Sandeep MS, Ravishankar MN (2014) The continuity of underperforming ICT projects in the public sector. Inf Manage 51(6):700–711
91. Sankar P, Schairer C, Coffin S (2003) Public mistrust: the unrecognized risk of the CDC smallpox vaccination program. Am J Bioeth 3(4):W22–W25
92. Scholz T, Pinney N (1995) Duty, fear, and tax compliance: the heuristic basis of citizenship behavior. Am J Political Sci 39:490–512
93. Schuller T, Baron S, Field J (2000) Social capital: a review and critique. In: Baron et al (eds) Social capital: critical perspectives. Oxford University Press, Oxford
94. Schyns P, Koop C (2010) Political distrust and social capital in Europe and the USA. Soc Indic Res 96(1):145–167

95. Sibley CG, Greaves LM, Satherley N, Wilson MS, Overall NC, Lee CHJ, Barlow FK (2020) Effects of the COVID-19 pandemic and nationwide lockdown on trust, attitudes toward government, and well-being. Am Psychol 75(5):618–630
96. Siegrist M, Zingg A (2014) The role of public trust during pandemics: implications for crisis communication. Eur Psychol 19:23–32
97. Smith RD, Keogh-Brown MR, Barnett T, Tait J (2009) The economy-wide impact of pandemic influenza on the UK: a computable general equilibrium modelling experiment. BMJ 339:b4571
98. Smith M (2010) Building institutional trust through e-government trustworthiness cues. Inf Technol People 23(3):222–246
99. Smith ML (2011) Limitations to building institutional trustworthiness through e-government: a comparative study of two e-services in Chile. J Inf Technol 26(1):78–93
100. Solomon RC, Flores F (2001) Building trust: in business, politics, relationships, and life. Oxford University Press, New York
101. Steyaert J (2000) Local government online and the role of the president: government shop versus the electronic community. Soc Sci Comput Rev 18(1):3–16
102. Teo TSH, Srivastava SC, Jiang L (2008) Trust and electronic government success: an empirical study. J Manag Inf Syst 25(3):99–132
103. Thomas W (1998) Maintaining and restoring public trust in government agencies and their employees. Adm Soc 30(2):166–193
104. Tolbert C, Mossberger K (2006) The effects of e-government on trust and confidence in government. Public Adm Rev 66(3):354–369
105. Toya H, Skidmore M (2014) Do Natural disasters enhance societal trust? Kyklos 67(2):255–279
106. Tschannen-Moran M, Hoy WK (2000) A multidisciplinary analysis of the nature, meaning, and measurement of trust. Rev Educ Res 70(4):547–593
107. Tyler R (1998) Trust and democratic government. In: Braithwaite V, Levi M, Trust and governance. Russell, Sage Foundation, New York
108. United Nations (2016) UN E-Government survey 2016. Online Dec 2017 at https://publicadministration.un.org/egovkb/en-us/reports/un-e-government-survey-2016
109. Van Bavel JJ, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, Willer R (2020) Using social and behavioural science to support covid-19 pandemic rsponse. Nat Hum Behav 4:460–471
110. Van Prooijen J-W, Van Dijk E (2014) When consequence size predicts belief in conspiracy theories: the moderating role of perspective taking. J Exp Soc Psychol 55:63–73
111. Vaughn A, Weldzius R (2021) Reshoring the global supply chains. Online Oct 2022 at http://www.ryanweldzius.com/uploads/1/0/7/2/107205599/reshoring_supply_chains_v2.pdf
112. Vinck P, Pham PN, Bindu KK, Bedford J, Nilles EJ (2019) Institutional trust and misinformation in the response to the 2018–19 Ebola outbreak in North Kivu, DR Congo: a population-based survey. Lancet Infect Dis 19(5):529–536
113. Warkentin D, Pavlou P, Rose G (2002) Encouraging citizen adoption of e-government by building trust. Electron Mark 12(3):157–162
114. Welch E, Hinnant C, Moon M (2005) Linking citizen satisfaction with e-government and trust in government. J Public Adm Res Theory 15(3):371–391
115. West DM (2005) Digital government: technology and public sector performance. Princeton University Press, Princeton
116. Whiteley P, Clarke HD, Sanders D, Stewart M (2016) Why do voters lose trust in governments? Public perceptions of government honesty and trustworthiness in Britain 2000–2013. Br J Politics Int Relat 18(1):234–254
117. Wilson MS, Rose C (2014) The role of paranoia in a dual-process motivationalmodel of conspiracy belief. In: Van Prooijen J-W, Van Lange PAM (eds), Power, politics, and paranoia. Cambridge University Press, New York, NY, pp 273–291

118. Yang K, LG. Anguelovl LG (2013) Trustworthiness of public service. In: Dwivedi YK, Shareef MA, Pandey SK, Kumar V (eds) Public administration reform: market demand from public organizations. Routledge, New York, NY
119. Zucker LG (1986) Production of trust: institutional sources of economic structure, 1840–1920. Organ Behav 8:53–111

The Demand for Big Data Skills in China

Xinyuan Lin, Wenjun Wang, and Fa-Hsiang Chang

Abstract The exponential growth of data has driven the rapid development of big data, reshaped the demands for skills in many industries, and created new jobs. Using detailed data on job requirements and posted wages from an online job site, we define big data skills as technical and general skills and estimate the demand and the value offered for big data skills. We find that 11.05% of job ads require big data technical skills, and organizations are willing to pay higher wages for labor with big data technical skills. The Scientific Research and Technology Services industry is willing to pay the highest value for big data technical skills. In all industries, the average posted wages of the job ads that require big data technical skills are higher than those that do not require big data technical skills, and the industries that have higher demands for big data technical skills are also providing higher posted wages in job ads that do not require big data technical skills.

Keywords Big data · Skill sets · Job advertisements

1 Introduction

During the past decade, all industries and public institutions have been producing new data at an unprecedented rate [1], which drives the evolution of big data. The definition of big data changes with the increase of data generated, from the earliest 'very large collections of data' [2, 3] to the widely used 3 V definition in 2001 [4], which refers to the characteristics in Volume (large volume of data), Variety (various types and sources of data), and Velocity (data quickly generated) of data, finally to today's 5 V—add Veracity (quality of data) and Value (value in the data) [5, 6], and the development of big data has reshaped the work content of many industries and created emerging new jobs [7]. Organizations have seen the potential of big data to improve productivity and decision-making effectiveness [8]. Thus, the demand

X. Lin · W. Wang · F.-H. Chang (✉)
Wenzhou-Kean University, Wenzhou, China
e-mail: fchang@kean.edu

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_57

for labor with big data skills is distributed in many industries. More and more jobs require big data skills, which leads to changes in the demand for labor skills. Despite the high demand, skilled labor is still short in supply [9, 10]. Therefore, organizations are willing to provide higher wages to attract talents with big data skills. However, there is minimum empirical evidence of the demand for big data skills and how much value organizations are willing to pay for the skills. This problem is becoming increasingly important since cultivating big data talents and digitalizing industries require understanding the current situation.

Therefore, this paper aims to enrich the fast-growing stream of studies that explore the impact of big data skills on the labor market. By analyzing job ads from one of the largest online job sites in China, we first develop a skill list for big data and provide some key facts about the demand for big data skills. We also explore the value organizations are willing to pay for big data skills at the whole market and industry levels.

This paper proceeds as follows. Section 2 reviews the previous literature concerning the definitions of big data skills. Section 3.1 describes our data, our definition of big data skills, and shows the summary statistics of the demand for big data skills in China. Section 3.2 represents the method used to explore the value organizations are willing to pay for big data skills. Section 4 provides evidence of the value offered for big data skills in the overall Chinese economy and various industries. Section 5 draws conclusions.

2 Skills

Most previous literature on big data skills [11–13] pays minimum attention to clarifying definitions of big data skills and explaining the source of demands for skills. Therefore, in this section, we review how the evolution of big data results in today's skill demands and clarify definitions of big data skills.

Today's big data is more about the ability to search, integrate, and extract value from large datasets than simply a concept [14]. Besides techniques and knowledge to deal with data, other widely required general skills are also essential for big data professionals to understand data and extract value from it accurately and work efficiently [11, 12, 15, 16]. Based on this, we divide the big data skills into:

(1) *Technical Skills*: the practices and IT tools widely used to generate value from big data, typically for professional fields including data analysis, business analysis, business intelligence, and data science.
(2) *General Skills*: personal traits and abilities broadly required by most occupations across the labor market. These traits and abilities improve interpersonal interaction efficiency and work performance.

2.1 Technical Skills

The evolution of big data has experienced four main stages [17, 18]. We attribute each skill field to the stage in which it grows most significantly to explain comprehensively.

Big Data 1.0 (1951–1996). Driven by the Digital Avalanche [19] and the World Wide Web [17], organizations began to use computers to collect and store data [20] to support operation and transaction information systems [17]. This developed the skills of:

1. *Database.* Use database languages to input, store, query, maintain, and analyze data to establish databases shared between departments and computer systems [21].
2. *Data Warehouse.* Use data warehouses to collect and clean transactional data, integrate data of different structures, and store and analyze historical data [22] to help online analysis and decision-making [23].
3. *Data Collection.* Find the appropriate data sources and bring data into the right environment for use [23], such as using Supervisory Control and Data Acquisition (SCADA) system to monitor industrial processes [24].

Big Data 2.0 (1997–2004). Web 1.0 motivated online commerce firms to analyze online user activities [25, 26], which combined with advances in graphics hardware [27] to develop data analysis and visualization. However, abuse of user data evoked the challenge of data security and privacy management [28–30]. This developed skills of:

4. *Data Analysis and Visualization.* Transform data into visual form and make data become understandable information to gain insight and knowledge [31].
5. *Management of Data Security and Privacy.* Prevent intentional inappropriate use of data, such as by completely anonymizing datasets to avoid privacy disclosure caused by personal identification [32].

Big Data 3.0 (2004–2015). The increase in user-generated content brought by Web 2.0 [17] drove the application of machine learning methods based on cloud services [25]. More and more organizations established large data centers and adopted top tier network technologies [33]. Distributed computing tools that are still popular in 2022, such as MapReduce [34] and Hadoop [35], emerged. Various cloud services were launched and widely used [36]. This developed the skills of:

6. *Distributed Computing.* Use related programming frameworks, packages, and other tools to serve the distributed system, a group of independently operating computing nodes that can interact with one another to complete a shared work [37, 38].
7. *Cloud Computing.* Use or construct a cloud for shared computing resources such as networks and servers [39] to accomplish work.

Big Data 4.0 (2015–). Internet of Things propelled using cloud and real-time stream data analysis to discover patterns when generating data [17] and allowed the extensive application of operational monitoring [17] and predictive analysis based on artificial intelligence and machine learning [3], which reinforced the need for the accuracy and reliability of the data collected [40]. This developed the skills of:

8. *Artificial Intelligence and Machine Learning.* Use external information to identify potential patterns by relying on machine learning [41]. Machine learning is a subset of Artificial Intelligence, which makes programs learn from data to improve algorithm automatically [42].
9. *Data Integration.* Merge disparate datasets with different syntaxes, schemas, and semantics to find model-establishment predictors [42, 43].
10. *Data Accuracy and Integrity Management.* Check and ensure the accuracy and integrity of data, such as deleting records with insignificant errors or outliers and looking for obvious defects in data and eliminating them [23].

The ten areas of skills introduced above are consistent with previous works concerning big data skills [11–13], except we try to avoid including programming to distinguish skills particularly for big data and unrelated IT skills.

2.2 General Skills

General skills, such as skills for communication with stakeholders, are personal traits that enhance interpersonal interactions and working performance [44]. On the one hand, previous literature shows that with the increase in demand for higher skills for labor, the focus of work is shifting toward general skills, and general skills are improved due to the existence of technical skills [45]. On the other hand, because technical skills evolve over time and gradually become obsolete [46], most employers require their employees to have general skills [11] to use technical skills better, to flexibly use general skills to adapt to different combinations of tasks [45], and to use general skills to replace fade technical skills [45] to contribute to the organization continuously [47].

Recent studies have proposed more specific arguments regarding how combining big data technical and general skills shapes the demand for general skills. Data analysts must know not only about data analysis and statistics but also about ethical and human behavior, business and organizational strategies, and how to understand and communicate with others regarding the insights extracted from data [48]. Since collaborating applications are increasingly used in the cloud [49] and the main task for database designing is to reflect demand from the physical world to data models [50], data engineers are required to have good cognitive and social skills to understand and analyze the need of end-users so that they can establish an effective database. Similarly, communication skills are essential for data scientists [8] and business analysts. A survey of employers in Malaysia finds that because researchers in organizations failed to understand and pay attention to the needs of data scientists, data scientists had to spend more time explaining their work than doing it, which results in the demand for literacy skills for data scientists [46]. Similarly, business analysts' primary work is transforming relevant insights into actual business impacts [51], which requires communication skills. Efficient data scientists also need a positive working attitude and strong executive ability to win people's hearts and promote effective interaction within the team [46].

Recent studies on big data skills provide evidence for the demand for general skills from various countries. Job ads containing the keyword 'big data' on Dice.com, an US job site, have been studied, which shows that communication, reporting, responsibility, and leadership are hot big data skills [51]. Vacancy ads in Canada, Australia, and the USA from Monster.com [13] are analyzed, which provide evidence that agile, planning, management, and consultancy as part of big data skills are popular in the labor market. As one of the first batches of papers in China that study big data skills and provide evidence on the salary of jobs with big data skills in China, our work enriches the existing studies in other countries that focus on big data skills.

Based on the above discussion, the general skills related to big data can be summarized in the following aspects: (1) cognitive, (2) social, (3) attitude, (4) literacy, and (5) executive. We propose our definition of big data skills in Sect. 3.1.

3 Data and Method

3.1 Data

Job vacancies on 51Job.com are our initial data source. We collect approximately 4.7 million job postings without duplicated ads. The specific dates of data collection are February 7, February 21, March 28, April 25, May 23, June 20, July 11, July 20, August 8, and August 22, 2022. The data incorporate detailed information about the job, such as job titles, requirements of education and experience, posted dates, posted wages, relevant details of firms including the name, the location, the industry, and job descriptions. We apply word segmentation through Python to separate words in the job descriptions, which assists our estimates of the demand for big data skills. The following subsections clarify how we define big data skills.

Technical Skills and General Skills of Big Data. Based on previous literature, we propose our definition of big data skills in Table 1. We divide big data skills into two major skills: technical skills and general skills. As discussed in Sect. 2, technical skills and general skills complement each other. Technical skills enhance the demand for general skills, and general skills smooth the effective utilization of technical skills.

Table 1 Definition and examples of keywords of big data skills

	Definitions	Examples of keywords (in Chinese)
Technical skills		
1. Data collection	Ability to find appropriate data sources and bring data into the right environment for use	提供数据, 数据来源, 数据录入, SCADA, 爬虫…
2. Data integration	Ability to bring the data into the appropriate environment for analysis, such as using technologies related to data communication, transmission, and cleaning	ETL, Datastage, 数据结构, Kettle, Web-Service Cognos, 数据通讯…
3. Data analysis and visualization	Abilities to analyze and visualize data to generate value and smooth communication of findings from data	数据处理, 数据管理,大数据, 数据分析…
4. Artificial Intelligence and machine learning	Ability to use IT technologies for Artificial Intelligence and machine learning to help analyze data deeply and extract value from data	Mahout, MLlib, automation, 人工智能, AI, 机器学习, 数智化…
5. Management of database and data warehouse	Ability to develop, use, and maintain database management systems and use IT technologies for data warehouses	database, 数据库, 数据中心, 数据仓库,数据库系统, 关系数据库…
6. Management of data accuracy and integrity	Ability to check and ensure the accuracy and integrity of data	测试数据, 保证数据, 数据完整性…
7. Management of data security and privacy	Abilities to ensure the safety and privacy of data users and data sources	数据安全, 数据备份, 信息安全…
8. Distributed computing	Ability to use IT technologies typically for supporting distributed computing	MapReduce, Zookeeper, Hadoop, YARN, Pig…
9. Cloud	Ability to use or construct cloud services to improve business capabilities and reduce data processing costs	Azure, AWS, Cloud, Vmware, OpenStack…
General skills		
1. Cognitive	Abilities to (1) create and develop original ideas or products, (2) overcome stress, (3) be careful and do tasks scrupulously	洞察力, 理解力, 理解能力,领悟力, 悟性…
2. Social	Ability to understand and communicate with others to build harmonious interpersonal relationships	协同工作, 协作关系, 协作性, 公关能力…
3. Attitude	Personalities of being (1) honest, (2) enthusiastic about work, (3) humble, and (4) responsible	积极主动, 主动, 主动性,干劲, 积极进取…

(continued)

Table 1 (continued)

	Definitions	Examples of keywords (in Chinese)
4. Literacy	Abilities to speak and write	口齿清楚, 口语…
5. Executive	Ability to react to changing working conditions, solve problems, and make decisions quickly and effectively	计划性, 整体规划, 周到, 能干, 实践性, 实战…

We set nine minor technical skills and five minor general skills, which collectively reflect the skills we discussed in Sect. 2. Keywords from job postings are distilled and classified into each minor category. Our keywords include three types of words:

1. Practices. Practices to extract benefits from big data, such as 'providing data ('提供数据' in Chinese),' which is included in 'Data collection' skills because it concerns getting data ready for future use.
2. Concepts. Concepts related to that skill, such as 'data source' ('数据来源' in Chinese), which is classified into 'Data Collection' skills.
3. IT Techniques. We match information techniques with skill categories based on the typical objectives that people use these techniques for. For techniques that can serve multiple objectives, we match them with the skill categories for which the techniques are developed. For instance, we include 'MapReduce' in 'Distributed Computing' because MapReduce is a programming model developed with the goal of realizing distributed algorithms effectively [52], despite MapReduce is also used for other functions such as machine learning and data visualization [53].

Since we only include the most popular words and phrases that describe practices, concepts, and techniques required in job ads, this list of keywords does not provide comprehensive skills. Other practices, concepts, and techniques may also be important but not included since they are not the most popular in our sample.

Demand for Big Data Skills. To estimate the requirement of big data skills in each posting, we use Python to detect whether a job posting requires at least one keyword of big data skills as listed in Table 1. For example, if 'MySQL, database' appears in the job description of a job ad, the technical skill of this job ad is tagged as one (we only detect whether one ad requires a specific skill rather than the extent it requires). Table 2 provides the summary statistics of big data skills, where 11.05% of job ads require technical skills, 31.78% of job ads require cognitive skills in general skills, and 15.82% of job ads require social skills in general skills.

Overview. We match small classes of occupations, locations, and industries in the postings to occupations in the Standard Occupation Classification (SOC), cities at prefecture level and above, and Chinese Industrial Classification, respectively. Table 2 shows summary statistics of our dataset. We match the data to 21 industrial classifications, 297 cities at the prefecture level or above, and 85 SOC minor occupations.

Table 2 Summary statistics

		Mean
Big data skills		
Technical		0.1105
General	– Cognitive	0.3178
General	– Social	0.1582
General	– Attitude	0.3507
General	– Literacy	0.0464
General	– Executive	0.3389
Experience		2.1883
Education		3.7004
Posted wage (in thousands, RMB)		8.5755
Number of SOC major occupations		22
Number of SOC minor occupations		85
Number of cities at prefecture level and above		297
Number of industrial classifications		21
Number of occupation-city cells		17,752
Number of postings, total		4,606,602

The average posted monthly wage is 8.58 thousand RMB. Zero is filled out if no experience or education requirements are specified in the posting. Experience ranges from zero to ten years. Education ranges from one to seven, representing different education levels.[1] We aggregate our data into 17,752 occupation-city cells as observations. The percentage of skill demand, the average experience and education, as well as the average posted wages within each occupation-city cell are measured to represent its requirements on the skills, experience, and education, as well as the average wages it offers.

3.2 Method

Big Data Skills Across Various Industries. As shown in Table 3, we calculate the percentage of postings requiring technical skills in each industry and rank industries in descending order. We present the top ten industries with the highest demand in Table 3 to show the distribution of demands in industries with the highest demand. Consistent with what is expected, the Software and IT Services industry shows the highest demand for big data technical skills, which reaches 23.32%. In the Scientific

[1] Education level: Not mentioned/required—0; Junior middle school and below—1; Technical secondary school—2; Senior middle school—3; Junior college—4; Bachelor's degree—5; Master's degree—6; Doctor's degree—7.

Research and Technology Services industry, 13.44% of postings require technical skills, followed by Wholesale and Retail, Synthesis, Public Administration, Social Security, and Social Organizations, whose demands are all around 10%.

Moreover, the average posted wage in all postings, the average posted wage in postings with technical skills requirement, and the average posted wage in postings without technical skills requirement in each industry are measured. The result manifests that industries with higher demand for the technical skills usually offer higher posted wages on average. Also, even within each industry, the average posted wage is higher for postings requiring technical skills than that for other postings. In the next section, we illustrate our regression model to explore the relationship between demands for big data skills and posted wages.

Regression Model. This paper explores whether the demand for big data skills impacts the posted wage. The regression model is established as follows:

$$\log(\text{Wage})_{\text{oc}} = \alpha + \beta_1 \text{BigData}_{\text{oc}} + \beta_2 \text{Exp}_{\text{oc}} + \beta_3 \text{Edu}_{\text{oc}} + \text{Controls} + \varepsilon_{\text{oc}} \quad (1)$$

where $\log(\text{Wage})_{\text{oc}}$ represents the log of the average monthly wage for an occupation o in a city c, and we regress it on the average requirement of big data skills $\left(\text{BigData}_{\text{oc}}\right)$ in the occupation-city cells. We control for the average years of experience as well as the average level of education requirements in the cells. Controls of occupation, city, and industry are added in succession. All regressions are weighted by the number of job postings in each occupation-city cell.

Table 3 Demand for technical skill of big data by industries

Industry	Tech. skills (%)	Avg. wage	Avg. wage—tech. skills	Avg. wage—non-tech. skills
Software and IT Services	0.2332	9.4758	11.3245	8.9136
Scientific Research and Technology Services	0.1344	9.5825	12.0676	9.1967
Wholesale and Retail	0.1077	7.2752	8.1776	7.1664
Synthesis	0.1066	8.9317	10.1312	8.7886
Public Administration, Social Security, and Social Organizations	0.1039	9.5506	11.3554	9.3413
Financial	0.0978	10.1707	11.6045	10.0153
Manufacturing	0.0964	8.5616	10.1563	8.3915
Electricity, Heat, Gas, and Water Production and Supply	0.0899	8.0520	10.0363	7.8559
Health and Social Work	0.0860	8.5466	9.5925	8.4481
Culture, Sports, and Entertainment	0.0856	7.3294	8.4909	7.2207

4 Results

4.1 Skill Requirements of Big Data and Posted Wages

Table 4 shows the regression results of Eq. (1) on our dataset. Column 1 indicates a model that incorporates the estimates of technical and general skills of big data, experience, and education variables. The coefficients of these variables are all statistically significant at the less than 1% level. The result shows that a 10% increase in the percentage of technical skill demands leads to an approximately 4.93% increase in the posted wage. In addition, the average posted wage rises by 3.44% if the requirements of education increase by one level and rises by 15.12% if the requirements of experience increase one more year.

Columns 2, 3, and 4 show the regression results with a set of controls: fixed effects of SOC minor occupations and cities at the prefectural level and above and the

Table 4 Average wages and skill requirements of big data

	Dependent variable: log mean wage			
	(1)	(2)	(3)	(4)
Technical	0.4925***	1.0316***	0.2808***	0.2224***
	(0.009)	(0.039)	(0.031)	(0.032)
Cognitive	0.2483***	0.4703***	0.0899***	0.1090***
	(0.018)	(0.022)	(0.018)	(0.018)
Social	0.4078***	0.0511	0.0328	0.0158
	(0.035)	(0.032)	(0.025)	(0.025)
Attitude	0.1492***	0.1823***	−0.2033***	− 0.1985***
	(0.022)	(0.021)	(0.017)	(0.017)
Literacy	− 0.7848***	0.7919***	0.1151***	0.0776*
	(0.057)	(0.056)	(0.043)	(0.043)
Executive	0.3978***	0.1886***	0.0626**	0.0741***
	(0.042)	(0.036)	(0.028)	(0.028)
Experience	0.1512***	0.1546***	0.0715***	0.0837***
	(0.002)	(0.004)	(0.003)	(0.003)
Education	0.0344***	0.1028***	0.0087**	0.0107**
	(0.002)	(0.004)	(0.004)	(0.004)
Occupation FE		X	X	X
City FE			X	X
Industry FE				X
Occupation-City cells	17,752	17,752	17,752	17,752
R^2	0.514	0.774	0.878	0.883
Adj. R^2	0.513	0.773	0.875	0.880

percentage of ads across different industrial classifications. These control variables are added to account for some factors that may influence the relationship between the requirement of big data skills and the posted wages. First, the posted wages are usually higher in cities with better economic growth though there is not much difference in the skill requirements. Second, for occupations and industries highly related to big data, some relevant skills may be taken for granted and not mentioned in the job descriptions. While adding these controls, the coefficient of big data skills is still statistically significant at the less than 1% level but slightly decreases to 0.2224.

The results in Table 4 reveal that the requirement of big data skills has strong explanatory power to the posted wage. Job ads that require big data skills tend to provide higher posted wages. Besides, results with control variables show that the positive impact of big data skills on the posted wages is not confined to high-tech occupations such as computer or mathematical jobs, well-developed regions, or prosperous industries.

4.2 Skill Requirements and Posted Wages Across Industries with Top Six Technical Skills Share

To further explore whether the differences in posted wages across industries are due to the skill requirements, we regress the demand for big data skills on the average posted wage of occupation-city cells across industries with the top six technical skill shares in Table 5. Experience, education, occupation, and city fix effects are also controlled. The results are shown in Table 5.

Column 1 reveals that in the Software and IT Services industry, a 10% increase in the percentage of technical skill requirements can add 0.505% to the average posted wages. Moreover, the industry of Scientific Research and Technology Services (column 2) provides the highest increase in the posted wage for the technical skills of big data. However, in the Synthesis industry, the coefficient of technical skill is not statistically significant, while general skills of cognitive and social are statistically significant at the 1% level and associated with 0.744% and 0.669% growth in the average posted wage, respectively, if 10% augmenting in their demand.

Table 5 Average wages and skill requirements across industries

	Dependent variable: log mean wage					
	Software and IT Services	Scientific Research and Technology Services	Wholesale and retail	Synthesis	Public Administration, Social Security, and Social Organizations	Financial
	(1)	(2)	(3)	(4)	(5)	(6)
Technical	0.0505*	0.2969***	0.1212***	0.0198	0.2619***	0.0913*
	(0.029)	(0.031)	(0.037)	(0.033)	(0.068)	(0.049)
Cognitive	0.0414**	0.0196	0.1058***	0.0744***	0.0399	0.0515*
	(0.021)	(0.020)	(0.020)	(0.019)	(0.039)	(0.027)
Social	0.0690**	– 0.0543**	0.0096	– 0.0669***	– 0.0198	0.2134***
	(0.028)	(0.027)	(0.025)	(0.024)	(0.045)	(0.031)
Attitude	– 0.0256	– 0.0925***	– 0.0976***	– 0.0010	– 0.0996***	– 0.0290
	(0.018)	(0.019)	(0.018)	(0.018)	(0.035)	(0.025)
Literacy	0.0738	0.0473	0.1604***	0.0176	– 0.0218	0.1284***
	(0.047)	(0.042)	(0.045)	(0.041)	(0.057)	(0.042)
Executive	– 0.0723**	0.0094	0.0226	– 0.0211	– 0.0009	0.3735***
	(0.031)	(0.026)	(0.029)	(0.026)	(0.057)	(0.043)
Experience	0.1376***	0.0375***	0.0697***	0.0900***	0.0633***	0.1034***

(continued)

Table 5 (continued)

	Dependent variable: log mean wage					
	Software and IT Services	Scientific Research and Technology Services	Wholesale and retail	Synthesis	Public Administration, Social Security, and Social Organizations	Financial
	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.006)
Education	0.0052	0.1024***	− 0.0033	0.0479***	0.1087***	− 0.0363***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.012)	(0.008)
Occupation FE	X	X	X	X	X	X
City FE	X	X	X	X	X	X
Occ.-city cells	8731	5479	6562	5065	1411	4120
R^2	0.862	0.826	0.796	0.827	0.805	0.788
Adj. R^2	0.856	0.813	0.783	0.814	0.760	0.766

5 Conclusion

In this paper, we use online job ads that include detailed job descriptions and wages to explore the demand for big data skills and the value organizations offer for these skills to show the impact of big data on the overall Chinese economy and various industries. We define big data skills and develop a list of the most frequently required skill keywords in job descriptions to show demands for big data skills in the labor market.

Our statistics show an 11.05% demand for technical skills, 31.78% demand for cognitive skills, and 15.82% demand for social skills in job ads at the level of the overall Chinese economy, suggesting that, on average, about one of each nine online job ads posted by organizations requires technical skills, which is relatively high. Social and cognitive skills are more popular than technical skills, indicating that, as general skills, they complement technical skills. The highest demand is in the Software and IT Services industry (23.32%), followed by the Scientific Research and Technology Services industry, implying that these industries rely more on big data skills. Further, our analysis of the average posted wage across industries indicates that organizations are offering higher wages for big data technical skills and the demand for technical skills results in higher average industry wages. Finally, at the overall market level, we find a posted wage increase of 4.93% for a 10% level increase in the demand for big data technical skills, and wage increases appear in both high-tech and low-tech occupations, indicating that employers offer a significant value for big data technical skills in a broad range of occupations and industries. The Scientific Research and Technology Services industry provides the highest value for big data technical skills, implying that the value of big data technical skills may rest in combining it with science.

Overall, our work enriches the literature studying the demand and value for big data skills in the labor market and fills the gap regarding the definition of big data skills and the source of the demand for big data skills. First, our work provides a clear definition of big data skills and reviews the sources of the demand for these skills. Second, our work shows a great demand for workers with big data skills, which provides further evidence for the global trend of transforming into a more data-driven era [54]. Third, by showing strong evidence that big data skills are considered valuable by employers, our work shows the potential return of big data from the perspective of employees.

However, as this paper is one of the first pieces of empirical evidence regarding the demand and value of big data skills, we may omit other factors impacting how big data shapes the demand for other skills. For instance, attitude in the general skills may be of high value but is often assumed by employers that the employee must have regardless of the wage offered, which may result in undiscovered patterns in the results. Future researchers may include other factors in their analysis to provide a more comprehensive examination. Additionally, the dynamic demand for big data skills over time is also worth studying. We leave these works for future researchers.

References

1. Mikalef P, Boura M, Lekakos G, Krogstie J (2019) Big data analytics and firm performance: findings from a mixed-method approach. J Bus Res 98:261–276. https://doi.org/10.1016/j.jbusres.2019.01.044
2. Weiss SM, Indurkhya N (1997) Predictive data mining: a practical guide. Morgan Kaufmann, San Francisco
3. Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manag 35(2):137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007
4. Laney D (2001) 3d data management: controlling data volume velocity and variety. META Gr Res Note 6(70):1
5. Terzo O, Ruiu P, Bucci E, Xhafa F (2013) Data as a service (Daas) for sharing and processing of large data collections in the cloud. In: 7th International conference on complex, intelligent, and software intensive systems. IEEE, Taichung, pp 475–480. https://doi.org/10.1109/CISIS.2013.87
6. Jain A, The 5 V's of big data. https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/
7. Mayer-Schönberger V, Cukier K (2014) Big data: a revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt, Boston
8. Chen H, Chiang RHL, Storey VC (2012) Business intelligence and analytics: from big data to big impact. MIS Q 36(4):1165–1188. https://doi.org/10.2307/41703503
9. Vijayarani S, Sharmila S (2016) Research in big data—an overview. Inform Eng Int J 4(3):1–20. https://doi.org/10.5121/ieij.2016.4301
10. Barro S, Davenport TH (2019) People and machines: partners in innovation. https://sloanreview.mit.edu/article/people-and-machines-partners-in-innovation/
11. Mikalef P, Giannakos MN, Pappas IO, Krogstie J (2018) The human side of big data: understanding the skills of the data scientist in education and industry. In: 2018 IEEE global engineering education conference (EDUCON). IEEE, Santa Cruz de Tenerife, pp 503–512. https://doi.org/10.1109/EDUCON.2018.8363273
12. Verma A, Yurov KM, Lane PL, Yurova YV (2019) An investigation of skill requirements for business and data analytics positions: a content analysis of job advertisements. J Educ Bus 94(4):243–250. https://doi.org/10.1080/08832323.2018.1520685
13. Debortoli S, Müller O, Vom Brocke J (2014) Comparing business intelligence and big data skills. Bus Inf Syst Eng 6:289–300 (2014). https://doi.org/10.1007/s12599-014-0344-2
14. Boyd D, Crawford K (2012) Critical questions for big data. Inf Commun Soc 15(5):662–679. https://doi.org/10.1080/1369118X.2012.678878
15. Bassellier G, Benbasat I (2004) Business competence of information technology professionals: conceptual development and influence on IT-business partnerships. MIS Q 28(4):673
16. Varian HR (2014) Big data: new tricks for econometrics. J Econ Perspect 28(2):3–28
17. Lee I (2017) Big data: dimensions, evolution, impacts, and challenges. Bus Horiz 60(3):293–303. https://doi.org/10.1016/j.bushor.2017.01.004
18. Barnes TJ (2013) Big data, little history. Dialogues Hum Geogr 3(3):297–302. https://doi.org/10.1177/2043820613514323
19. Hacking I (1991) How should we do the history of statistics? In: Burchell G, Gordon C, Miller P (eds) The Foucault effect: studies in governmentality: with two lectures by and an interview with Michel Foucault. University of Chicago Press, Chicago, pp 181–195
20. Tukey JW (1962) The future of data analysis. Annu Math Stat 33(1):1–67
21. Chamberlin DD (1976) Relational data-base management systems. ACM Comput Surv 8(1):43–66. https://doi.org/10.1145/356662.356665
22. Inmon WH (1996) The data warehouse and data mining. Commun ACM 39(11):49–50.https://doi.org/10.1145/240455.240470
23. Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) From data mining to knowledge discovery in databases. AI Mag 17(3):37. https://doi.org/10.1609/aimag.v17i3.1230
24. Daneels A, Salter W (1999) What is SCADA? In: International conference on accelerator and large experimental physics control systems. JACoW, Trieste, pp 339–343

25. Beer D (2016) How should we do the history of big data? Big Data Soc 1:1–10. https://doi.org/10.1177/2053951716646135
26. Davenport TH (2006) Competing on analytics. Harv Bus Rev 84(1):98–107. https://hbr.org/2006/01/competing-on-analytics
27. Van Wijk JJ (2006) Views on visualization. IEEE Trans Vis Comput Graph 12(4):421–432. https://doi.org/10.1109/TVCG.2006.80
28. Brophy P, Halpin E (1999) Through the net to freedom: information, the internet and human rights. J Inf Sci 25(5):351–364. https://doi.org/10.1177/016555159902500502
29. Seltzer W, Anderson M (2022) The dark side of numbers: the role of population data systems in human rights abuses. Soc Res 68(2):481–513. https://www.jstor.org/stable/40971467
30. Miller S (2014) Collaborative approaches needed to close the big data skills gap. J Organ Des 3(1):26. https://doi.org/10.7146/jod.9823
31. Li Q (2020) Overview of data visualization. In: Li Q (ed) Embodying data: Chinese aesthetics, interactive visualization and gaming technologies. Springer, Singapore, pp 17–47
32. Narayanan A, Shmatikov V (2009) De-anonymizing social networks. In: 30th IEEE symposium on security and privacy. IEEE, Oakland, pp 173–187
33. Anonymous (2008) Community cleverness required. Nature 455:1. https://doi.org/10.1038/455001a
34. Gillick D, Faria A, DeNero J (2006) Mapreduce: distributed computing for machine learning. Berkley 12
35. Shvachko K, Kuang H, Radia S, Chansler R (2010) The hadoop distributed file system. In: IEEE 26th symposium on mass storage systems and technologies (MSST). IEEE, Nevada, pp 1–10. https://doi.org/10.1109/MSST.2010.5496972
36. Agarwal A, Siddharth S, Bansal P (2016) Evolution of cloud computing and related security concerns. In: 2016 Symposium on colossal data analysis and networking (CDAN). IEEE, Indore, pp 1–9. https://doi.org/10.1109/CDAN.2016.7570920
37. Garcia-Molina H (1982) Elections in a distributed computing system. IEEE Trans Comput 31(1):48–59. https://doi.org/10.1109/TC.1982.1675885
38. Rothnie JB, Goodman N (1977) A survey of research and development in distributed database management. In: Housel BC, Merten AG (eds) Proceedings of the third International Conference on very Large Data Bases. VLDB, vol 3. VLDB Endowment, Tokyo, pp 48–62. https://dl.acm.org/doi/abs/10.5555/1286580.1286585
39. Jadeja Y, Modi K (2012) Cloud computing—concepts, architecture and challenges. In: 2012 International conference on computing, electronics and electrical technologies (ICCEET). IEEE, Nagercoil, pp 877–880
40. Bragazzi NL, Dai H, Damiani G, Behzadifar M, Martini M, Wu J (2020) How big data and artificial intelligence can help better manage the covid-19 pandemic. Int J Environ Res Public Health 17(9):3176. https://doi.org/10.3390/ijerph17093176
41. Kaplan A, Haenlein M (2019) Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Bus Horiz 62(1):15–25. https://doi.org/10.1016/j.bushor.2018.08.004
42. Guo Y, Zhang Y, Lyu T, Prosperi M, Wang F, Xu H, Bian J (2021) The application of artificial intelligence and data integration in covid-19 studies: a scoping review. J AMIA 28(9):2050–2067. https://doi.org/10.1093/jamia/ocab098
43. Lenzerini M (2002) Data integration: a theoretical perspective. In: Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on principles of database systems, PODS '02. Association for Computing Machinery, New York, pp 233–246. https://doi.org/10.1145/543613.543644
44. Cacciolatti L, Lee SH, Molinero CM (2017) Clashing institutional interests in skills between government and industry: an analysis of demand for technical and soft skills of graduates in the UK. Techol Forecast Soc Change 119:139–153. http://dx.doi.org/10.1016/j.techfore.2017.03.024
45. Grugulis I, Vincent S (2009) Whose skill is it anyway? 'Soft' skills and polarization. Work Employ Soc 23(4):597–615. https://doi.org/10.1177/0950017009344862

46. Saari A, Rasul MS, Mohamad Yasin R, Abdul Rauf RA, Mohamed Ashari ZH, Pranita D (2021) Skills sets for workforce in the 4th industrial revolution: expectation from authorities and industrial players. J Tech Educ Train 13(2):1–9. https://doi.org/10.30880/jtet.2021.13.02.001
47. Schmeelk S, Dragos D (2020) NICE framework special issue: investigating framework adoption, adaptation, or extension. https://www.researchgate.net/publication/352208372
48. Malaysian Ministry of Higher Education. https://www.researchgate.net/publication/330506612
49. Andres B, Poler R, Sanchis R (2021) A data model for collaborative manufacturing environments. Comput Ind 126:103398 (2021). https://doi.org/10.1016/j.compind.2021.103398
50. Zhou B, Wang S, Xi L (2005) Data model design for manufacturing execution system. J Manuf Technol Manage 16(8):909–935 (2005). https://doi.org/10.1108/17410380510627889
51. De Mauro A, Greco M, Grimaldi M, Ritala P (2018) Human resources for big data professions: a systematic classification of job roles and required skill sets. Inf Process Manage 54(5):807–817. https://doi.org/10.1016/j.ipm.2017.05.004
52. De Mauro A, Greco M, Grimaldi M (2015) What is big data? A consensual definition and a review of key research topics. AIP Conf Proc 1644(1):97. https://doi.org/10.1063/1.4907823
53. Dean J, Ghemawat S (2008) Mapreduce: simplified data processing on large clusters. Commun ACM 51(1):107–113. https://doi.org/10.1145/1327452.1327492
54. Perera-Tallo F (2017) Growing income inequality due to biased technological change. J Macroecon 52:23–38. https://doi.org/10.1016/j.jmacro.2017.02.002

Set-Membership Filtering for 2-D Systems with State Constraints Under the FlexRay Protocol

Meiyu Li and Jinling Liang

Abstract The states of some practical dynamical systems meet certain constraints, which have to be considered when estimating the corresponding states. This chapter studies the issue of set-membership filtering (SMF) with a state-constrained two-dimensional system. The signal transmission is adjusted using the FlexRay protocol to lessen communication burden and increase data scheduling flexibility. We propose a recursive algorithm, utilizing the set-membership technique and induction, for finding a set of optimal ellipsoids containing the actual states at every position in the presence of state equality constraint, the FlexRay protocol, and non-Gaussian noises. Numerical results illustrate effectiveness of the proposed state equality-constrained SMF design scheme.

Keywords Two-dimensional systems · Set-membership filtering · State constraint · FlexRay protocol

1 Introduction

The advancement of modern industry, as well as the demand for multivariable analysis, has heightened interest in two-dimensional (2-D) systems with state variables propagated in two separate orientations. A special property of 2-D models is two-way propagation, as opposed to the typical one-dimensional (1-D) models that develop unidirectionally. In the pioneering works [1], this property has been exploited based on the Roesser model to investigate the multi-dimensional iterative circuits and the linear image processing. Subsequently, as a more extensive model, the classical first-type and second-type Fornasini-Marchesini models are proposed and their state space implementation theories are established [2]. Based on them, the research on

M. Li (✉) · J. Liang
Southeast University, Nanjing 210096, China
e-mail: limeiyuzc99@126.com

J. Liang
e-mail: jinlliang@seu.edu.cn

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_58

2-D systems has developed vigorously in recent decades. As of now, a large number of research accomplishments have been made for 2-D systems in the areas of industrial control, images filtering, communication fault detection, and so on [3–5].

Filtering is an important measure for suppressing interference because it would remove various interferential signals from contaminated measurement information. Thereinto, set-membership filtering (SMF) has gained an increasing academic/industrial interest because it is dependent merely on the information of the hard boundaries of the system states and the external disturbances (i.e., the process and measurement disturbances). The SMF aims to locate an area in the state space where the unknown-but-actual state vectors belong to. As a result, rather than providing the most viable state in which certain optimality such as the Kalman filtering or the H_∞ filtering, the SMF issue tries to identify the smallest feasible state estimation set. To date, majority of the SMF researches are concentrated on 1-D systems, and the recent interesting ones can be found in [6, 7]. While besides the pioneering works are given in [8, 9], the relating SMF results for 2-D systems are relatively lacking, which is just the major research motivation of this work.

In some practical systems, the system variables need to meet certain special constraints, which include, but are not limited to, the basic laws of physics, the kinematic or geometric factors of the systems, and the mathematical descriptions of certain state vectors. Their engineering applications include ground target tracking (e.g., traffic rules, physical road network constraints), maritime navigation (e.g., coastline), etc. [10]. These applications further infer that reasonable processing/utilization of these constraints information, which usually can be modeled as equality (or inequality) constraints, would improve the estimation accuracy. For some recent relating developments, one might refer to [11, 12]. As far as the authors know, the SMF under state constraint for 2-D systems has not been explored, which is the second motivation for this present paper.

On the other hand, with the advancement of network technology in recent years, more and more systems tend to be networked. Frequent data transmission will unavoidably result in different types of networked phenomena such as network congestion and delays because of the limited bandwidth of the communication channels [13–16]. Data congestion is one of them that significantly affects how well the networked complex systems being handled perform. Introducing communication protocols is a practical technique to reduce this issue. The FlexRay protocol (FRP), a mixed data transmission, combines the benefits of time-triggered and event-triggered mechanisms and can flexibly select the data transfer modes in predefined time windows. This protocol's primary objective is to address the present high bandwidth, dependability, and certainty requirements, which has been successfully applied in many practical industrial fields [17–19]. It is notable that for the SMF of 2-D systems, the corresponding findings are scattered (if not none), let alone taking the state equality constraints into account simultaneously.

Inspired from the above analysis, this paper is devoted to studying the SMF problem for 2-D systems with state constraints under FRP. Based on the inductive method, the existence condition is obtained for designing an appropriate set-membership filter, and then an effective algorithm is proposed for obtaining the optimal ellipsoids containing the states of the original 2-D system. This chapter's primary contributions are: (1) investigating a kind of state-restricted 2-D systems' SMF issue under the FRP; (2) designing a recursive filter with expected filtering performance; (3) proposing a recursive algorithm that can be implemented online to obtain the optimized ellipsoids containing the real states of the addressed 2-D system.

2 Problem Formulation

Consider the 2-D system shown below:

$$\begin{cases} x_{j+1,k+1} = A^{(1)}_{j+1,k}x_{j+1,k} + A^{(2)}_{j,k+1}x_{j,k+1} + B^{(1)}_{j+1,k}\omega_{j+1,k} + B^{(2)}_{j,k+1}\omega_{j,k+1} \\ y_{j,k} = C_{j,k}x_{j,k} + D_{j,k}v_{j,k}, \end{cases} \tag{1}$$

where $j, k \in \mathbb{N}$ ($\mathbb{N} := \{0, 1, \ldots\}$), $x_{j,k} \in \mathbb{R}^{n_x}$ and $y_{j,k} \in \mathbb{R}^{n_y}$ are the state vector and the ideal output, respectively. $A^{(1)}_{j,k}$, $A^{(2)}_{j,k}$, $B^{(1)}_{j,k}$, $B^{(2)}_{j+1,k}$, $C_{j,k}$, and $D_{j,k}$ are known shift-varying matrices. $\omega_{j,k} \in \mathbb{R}^{n_\omega}$ and $v_{j,k} \in \mathbb{R}^{n_v}$ are the unknown external disturbances satisfying

$$\omega^T_{j,k}W^{-1}_{j,k}\omega_{j,k} \le 1, \quad v^T_{j,k}V^{-1}_{j,k}v_{j,k} \le 1 \tag{2}$$

with matrices $W_{j,k} > 0$ and $V_{j,k} > 0$. In addition, suppose that $x_{j,k}$ is required to satisfy the following equality constraint:

$$S_{j,k}x_{j,k} = s_{j,k}, \tag{3}$$

where matrix $S_{j,k}$ and vector $s_{j,k} \in \mathbb{R}^{n_s}$ are known with $1 \le n_s \le n_x$. The initial conditions with regard to (1) are $x_{j,k} = c_{j,k}$ when $j = 0$, $k \in \lfloor 0, \kappa_1 \rfloor$; $x_{j,k} = d_{j,k}$ when $j \in \lfloor 0, \kappa_2 \rfloor$, $k = 0$; and $x_{j,k} = 0$ otherwise; where $\lfloor 0, \kappa_1 \rfloor$ indicates the set $\{0, 1, \ldots, \kappa_1\}$, $c_{j,k}$ and $d_{j,k}$ are known vectors satisfying $c_{0,0} = d_{0,0}$.

In this paper, the output signals are transmitted to the filter via a communal network. In order to effectively reduce the network load, a hybrid communication protocol (i.e., a FRP) is utilized. As is well known, the FRP is a hybrid protocol which consists of a static segment where the periodically transmitted messages are critical to time and a dynamic segment, where the sporadically transmitted messages are triggered by events. Just as in [17, 19], we implement the round-robin protocol (RRP) and the try-once-discard protocol (TODP) in the static and the dynamic scheduling parts of the FlexRay protocol, respectively.

Specifically, the n_y sensors measuring each entry of the output are labeled by $1, 2, \ldots, n_y$, which are divided into two parts, the first $\hbar$ ($\hbar \ge 1$) belongs to set $\mathcal{S}_1 := \{1, 2, \ldots, \hbar\}$ and the remaining belongs to set $\mathcal{S}_2 := \{\hbar + 1, \hbar + 2, \ldots, n_y\}$. Without sacrificing generality and taking into account the various real-time needs, imagine the nodes that are part of $\mathcal{S}_1$ adopting the RRP and the rest belonging to $\mathcal{S}_2$

being scheduled by the TODP. Next, detail descriptions of the transmission rules as below.

• RRP:

$$\xi_{j,k} = \mathrm{mod}(j + k - 1, \hbar) + 1 \tag{4}$$

where $\xi_{j,k} \in \mathcal{S}_1$ represents the index of the sensor node which is chosen the right to transmit its measurement through the shared channels at instant (j, k).

• TODP:

$$\sigma_{j,k} = \arg \max_{l=\hbar+1,\ldots,n_y} \{\tilde{y}^{(l)}_{j,k} \Omega_l \tilde{y}^{(l)*}_{j,k}\} \tag{5}$$

where $\sigma_{j,k} \in \mathcal{S}_2$ denotes the mode index that has permission to use the communal network at point (j, k), $\tilde{y}^{(l)}_{j,k} = y^{(l)}_{j,k} - y^{(l)*}_{j,k}$ with $y^{(l)}_{j,k}$ being the l-th entry in vector $y_{j,k}$, $y^{(l)*}_{j,k}$ being the last signal transmitted by node l before instant (j, k), and Ω_l $(l \in \mathcal{S}_2)$ are given positive weighting coefficients.

Remark 1 The order of the double indices is defined as : $\forall j_1, k_1, j_2, k_2 \in \mathbb{N}$,

$$\begin{aligned}(j_1, k_1) < (j_2, k_2) \Leftrightarrow (j_1, k_1) \in \{(j, k) \in \mathbb{N} \times \mathbb{N} \mid j = j_2, k_1 < k_2\} \\ \cup\{(j, k) \in \mathbb{N} \times \mathbb{N} \mid j_1 < j_2\}.\end{aligned}$$

For convenience of expression, we denote the column vector $y^{[1]}_{j,k}$ composed of the first $\hbar$ components of $y_{j,k}$ and that the rest $n_y - \hbar$ terms constitute $y^{[2]}_{j,k} \in \mathbb{R}^{n_y-\hbar}$. The following network transmission, $\bar{y}_{j,k}$ represents the actual received output data, in which $\bar{y}^{[1]}_{j,k}$ and $\bar{y}^{[2]}_{j,k}$ represent the relevant parts scheduled by the abovementioned two different protocols, respectively. Combining the idea of zero-order hold with the characteristic of the FRP. It is noteworthy that $\bar{y}^{[1]}_{j,k}$ and $\bar{y}^{[2]}_{j,k}$ are set as 0 when $k < 0$.

Then, the real output $\bar{y}_{j,k}$ transmitted to the filter is represented as

$$\begin{aligned}\bar{y}_{j,k} &= \mathcal{I}_1 \bar{y}^{[1]}_{j,k} + \mathcal{I}_2 \bar{y}^{[2]}_{j,k} \\ &= \mathcal{I}_1 \Phi_1(\xi_{j,k})(C^{[1]}_{j,k} x_{j,k} + D^{[1]}_{j,k} v_{j,k}) + \mathcal{I}_1(I - \Phi_1(\xi_{j,k}))\bar{y}^{[1]}_{j,k-1} \\ &\quad + \mathcal{I}_2 \Phi_2(\sigma(j, k))(C^{[2]}_{j,k} x_{j,k} + D^{[2]}_{j,k} v_{j,k}) + \mathcal{I}_2(I - \Phi_2(\sigma_{j,k}))\bar{y}^{[2]}_{j,k-1}\end{aligned} \tag{6}$$

where $\Phi_1(\xi_{j,k}) = \mathrm{diag}_{1\le s\le\hbar}\{\delta(s - \xi_{j,k})\}$, $\Phi_2(\sigma_{j,k}) = \mathrm{diag}_{\hbar+1\le t\le n_y}\{\delta(t - \sigma_{j,k})\}$, $C^{[1]}_{j,k} = \mathrm{col}_{1\le s\le\hbar}\{C^{(s)}_{j,k}\}$, $C^{[2]}_{j,k} = \mathrm{col}_{\hbar+1\le t\le n_y}\{C^{(t)}_{j,k}\}$, $D^{[1]}_{j,k} = \mathrm{col}_{1\le s\le\hbar}\{D^{(s)}_{j,k}\}$, $D^{[2]}_{j,k} = \mathrm{col}_{\hbar+1\le t\le n_y}\{D^{(t)}_{j,k}\}$, $\mathcal{I}_1 := [I_\hbar, 0_{\hbar\times(n_y-\hbar)}]^T, \mathcal{I}_2 := [0_{(n_y-\hbar)\times\hbar}, \ I_{n_y-\hbar}]^T$; in which $I_\hbar$ is the $\hbar$-dimensional identity matrix, $0_{\hbar\times(n_y-\hbar)}$ means the $\hbar \times (n_y - \hbar)$ dimensional zero matrix, $C^{(l)}_{j,k}$ and $D^{(l)}_{j,k}$ with $l = 1, 2, \ldots, n_y$ are the l-th row vectors of matrices $C_{j,k}$ and $D_{j,k}$, respectively. Denote $\bar{x}_{j,k} = [x^T_{j,k}, (\bar{y}^{[1]}_{j,k-1})^T, (\bar{y}^{[2]}_{j,k})^T]^T$ and $\bar{\omega}_{j,k} = [\omega^T_{j,k}, v^T_{j,k}]^T$, then the 2-D system (1) with FRP can be expressed as

$$\begin{aligned}\bar{x}_{j,k} &= \bar{A}^{(1)}_{j,k-1}\bar{x}_{j,k-1} + \bar{A}^{(2)}_{j-1,k}\bar{x}_{j-1,k} + \bar{B}^{(1)}_{j,k-1}\bar{\omega}_{j,k-1} + \bar{B}^{(2)}_{j-1,k}\bar{\omega}_{j-1,k} \\ \bar{y}_{j,k} &= \bar{C}_{j,k}\bar{x}_{j,k} + \bar{D}_{j,k}\bar{\omega}_{j,k}, \quad i,\, j \in \mathbb{N}_+ \end{aligned} \tag{7}$$

where $\bar{A}^{(2)}_{j,k} = \mathrm{diag}\{A^{(2)}_{j,k}, 0, 0\}$, $\bar{D}_{j,k} = [0,\ \mathcal{I}_1\Phi_1(\xi_{j,k})D^{[1]}_{j,k} + \mathcal{I}_2\Phi_2(\sigma_{j,k})D^{[2]}_{j,k}]$,

$$\begin{aligned}\bar{A}^{(1)}_{j,k} &= \begin{bmatrix} A^{(1)}_{j,k} & 0 & 0 \\ \Phi_1(\xi_{j,k})C^{[1]}_{j,k} & I - \Phi_1(\xi_{j,k}) & 0 \\ \Phi_2(\sigma_{j,k})C^{[2]}_{j,k} & 0 & I - \Phi_2(\sigma_{j,k}) \end{bmatrix}, \\ \bar{B}^{(1)}_{j,k} &= \begin{bmatrix} B^{(1)}_{j,k} & 0 \\ 0 & \Phi_1(\xi_{j,k})D^{[1]}_{j,k} \\ 0 & 0 \end{bmatrix}, \quad \bar{B}^{(2)}_{j,k} = \begin{bmatrix} B^{(2)}_{j,k} & 0 \\ 0 & 0 \\ 0 & \Phi_2(\sigma_{j,k})D^{[2]}_{j,k} \end{bmatrix}, \\ \bar{C}_{j,k} &= [\mathcal{I}_1\Phi_1(\xi_{j,k})C^{[1]}_{j,k} + \mathcal{I}_2\Phi_2(\sigma_{j,k})C^{[2]}_{j,k}, \\ &\qquad \mathcal{I}_1(I - \Phi_1(\xi_{j,k})),\ \mathcal{I}_2(I - \Phi_2(\sigma(j,k)))].\end{aligned}$$

For system (7), we construct the following recursive filter:

$$\begin{aligned}\hat{x}_{j,k} &= \bar{A}^{(1)}_{j,k-1}\hat{x}_{j,k-1} + \bar{A}^{(2)}_{j-1,k}\hat{x}_{j-1,k} + K^{(1)}_{j,k-1}[\bar{y}_{j,k-1} - \bar{C}_{j,k-1}\hat{x}_{j,k-1}] \\ &\quad + K^{(2)}_{j-1,k}[\bar{y}_{j-1,k} - \bar{C}_{j-1,k}\hat{x}_{j-1,k}], \quad i,\, j \in \mathbb{N}_+ \end{aligned} \tag{8}$$

where $j, k \in \mathbb{N}_+ := \mathbb{N}\backslash\{0\}$, $\hat{x}_{j,k} \in \mathbb{R}^{n_x+n_y}$ is an estimate of $\bar{x}_{j,k}$, matrices $K^{(1)}_{j,k}$ and $K^{(2)}_{j,k}$ are the filter gains to be solved. Set $\hat{x}_{j,0} = u_{j,k}$, $\hat{x}_{0,k} = g_{j,k}$ for $j, k \in \mathbb{N}$ with $u_{0,0} = g_{0,0}$ as the initial states of filter (8). We denote $e_{j,k} = \bar{x}_{j,k} - \hat{x}_{j,k}$ as the filtering error.

Assumption 1 The filtering error system's initial boundary states satisfy

$$e^T_{j,0}R^{-1}_{j,0}e_{j,0} \le 1, \quad e^T_{0,k}R^{-1}_{0,k}e_{0,k} \le 1$$

for any $j, k \in \mathbb{N}$, where matrices $R_{j,0} > 0$ and $R_{0,k} > 0$ are known.

This chapter's purpose is to determine a series of ellipsoids for the given measurement information $\bar{y}_{j,k}$, the unknown-but-bounded (UBB) disturbances $\omega_{j,k}$ and $v_{j,k}$, and the state constraint (3). To put this in another way, we are looking for a sequence of matrices $R_{j,k} > 0$, $K^{(1)}_{j,k}$, and $K^{(2)}_{j,k}$ with $j, k \in \mathbb{N}$ such that, under Assumption 1 the filtering error $e_{j,k}$ always satisfies

$$e^T_{j,k}R^{-1}_{j,k}e_{j,k} \le 1, \quad \forall j, k \in \mathbb{N} \tag{9}$$

subject to constraint (3). Then, to obtain an optimal ellipsoid, matrix $R_{j,k}$ is minimized in the sense of trace at each point.

3 Main Results

This section examines the SMF issue for the discrete shift-varying system (1) with state constraint (3) under the FlexRay protocol. First, sufficient conditions are presented for calculating the state estimation ellipsoid. Then, by properly designing the filter gains, a recursive algorithm is developed to minimize $R_{j,k}$ (in the sense of trace). In order to proceed, the below helpful lemma is still required.

Lemma 1 *([20]) Let $\psi \in \mathbb{R}^q$, $P = P^T \in \mathbb{R}^{q\times q}$, and $H \in \mathbb{R}^{p\times q}$ with* $\operatorname{rank}(H) = r < q$. *In such cases, the subsequent statements are parallel: 1) $\psi^T P\psi \le 0$ for all $\psi \ne 0$ satisfying $H\psi = 0$; 2) $(H^\perp)^T PH^\perp \le 0$; 3) $\exists\, \varsigma \in \mathbb{R}$ s.t. $P - \varsigma H^T H \le 0$; 4) $\exists\, E \in \mathbb{R}^{p\times q}$ s.t. $P + E^T H + H^T E \le 0$; where $H^\perp$ is a right orthogonal complement of matrix H.*

Theorem 1 *Consider the 2-D DSVS (1) with state constraint (3), the FRP given by (4)–(5) and the corresponding shift-varying filter (8). Let the initial matrices $R_{j,0}$ and $R_{0,k}$ with $j, k \in \mathbb{N}$ be given. The filtering error $e_{j+1,k+1}$ always satisfies constraint (9) for all $j, k \in \mathbb{N}$ under Assumption 1 if there exist matrices $R_{j+1,k+1} > 0$, $K^{(1)}_{j+1,k}$, $K^{(2)}_{j,k+1}$, $N_{j,k}$, scalars $\alpha^{(r)}_{j,k}$ $(r = 1, 2, 3, 4)$ such that*

$$\begin{bmatrix} -R_{j+1,k+1} & \tilde{\Psi}_{j,k} \\ \tilde{\Psi}^T_{j,k} & -\Xi_{j,k} \end{bmatrix} \le 0 \tag{10}$$

holds for all $j, k \in \mathbb{N}$, where

$$\begin{aligned}
&\tilde{\Psi}_{j,k} = \big[0, \tilde{\Psi}^{(2)}_{j,k}, \tilde{\Psi}^{(3)}_{j,k}, \tilde{\Psi}^{(4)}_{j,k}, \tilde{\Psi}^{(5)}_{j,k}\big], \quad \tilde{\Psi}^{(2)}_{j,k} = \big(\bar{A}^{(1)}_{j+1,k} - K^{(1)}_{j+1,k}\bar{C}_{j+1,k}\big)L_{j+1,k},\\
&\tilde{\Psi}^{(3)}_{j,k} = \big(\bar{A}^{(2)}_{j,k+1} - {}^{(2)}_{j,k+1}\bar{C}_{j,k+1}\big)L_{j,k+1}, \quad \tilde{\Psi}^{(4)}_{j,k} = \bar{B}^{(1)}_{j+1,k} - K^{(1)}_{j+1,k}\bar{D}_{j+1,k},\\
&\tilde{\Psi}^{(5)}_{j,k} = \bar{B}^{(2)}_{j,k+1} - K^{(2)}_{j,k+1}\bar{D}_{j,k+1}, \quad \Xi_{j,k} = \Upsilon_{j,k} - \operatorname{sym}\{N^T_{j,k}(\Pi^{(1)}_{j,k} + \Pi^{(2)}_{j,k})\},\\
&\Upsilon_{j,k} = \Psi^{(0)} + \sum_{r=1}^{2}\alpha^{(r)}_{j,k}\Upsilon^{(r)} + \sum_{r=3}^{4}\alpha^{(r)}_{j,k}\tilde{\Upsilon}^{(r)}_{j,k}, \ \Psi^{(0)} = \operatorname{diag}\{1, 0, 0, 0, 0\},\\
&\Upsilon^{(1)} = \operatorname{diag}\{-1, I, 0, 0, 0\}, \quad \Upsilon^{(2)} = \operatorname{diag}\{-1, 0, I, 0, 0\},\\
&\tilde{\Upsilon}^{(3)}_{j,k} = \operatorname{diag}\{-2, 0, 0, Q^{-1}_{j+1,k}, 0\}, \quad \tilde{\Upsilon}^{(4)}_{j,k} = \operatorname{diag}\{-2, 0, 0, 0, Q^{-1}_{j,k+1}\},\\
&\Pi^{(1)}_{j,k} = [\bar{S}_{j+1,k}\hat{x}_{j+1,k} - s_{j+1,k}, \bar{S}_{j+1,k}L_{j+1,k}, 0, 0, 0],\\
&\Pi^{(2)}_{j,k} = [\bar{S}_{j,k+1}\hat{x}_{j,k+1} - s_{j,k+1}, 0, \bar{S}_{j,k+1}L_{j,k+1}, 0, 0],
\end{aligned}$$

in which $\operatorname{sym}\{G\} := G + G^T$, $Q^{-1}_{j,k} = \operatorname{diag}\{W^{-1}_{j,k}, V^{-1}_{j,k}\}$, $\bar{S}_{j,k} = [S_{j,k}, 0, 0]$, *and $L(j, k)$ is the Cholesky factorisation of $R_{j,k}$, that is, $R_{j,k} = L_{j,k}L^T_{j,k}$.*

Proof Firstly, according to Assumption 1, for all $(j, k) \in \{(j_0, k_0) : j_0, k_0 \in \mathbb{N}, j_0 + k_0 = 1\}$, the filtering error obviously satisfies project object (9).

From (7)–(8), the one-step-forward estimation error could be expressed as

$$\begin{aligned} e_{j+1,k+1} = \ & \bar{A}^{(1)}_{j+1,k} e_{j+1,k} + \bar{A}^{(2)}_{j,k+1} e_{j,k+1} \\ & + \bar{B}^{(1)}_{j+1,k} \bar{\omega}_{j+1,k} + \bar{B}^{(2)}_{j,k+1} e_{j,k+1} \bar{\omega}_{j,k+1} \\ & - K^{(1)}_{j+1,k} [\bar{C}_{j+1,k} e_{j+1,k} + \bar{D}_{j+1,k} \bar{\omega}(j+1,k)] \\ & - K^{(2)}_{j,k+1} e_{j,k+1} [\bar{C}_{j,k+1} e_{j,k+1} + \bar{D}_{j,k+1} \bar{\omega}_{j,k+1}]. \end{aligned} \tag{11}$$

Secondly, assume that $e_{j+1,k}$ and $e_{j,k+1}$ satisfy the projective object (9). Then there exist vectors $z_{j+1,k}$ and $z_{j,k+1}$ satisfying $\| z_{j+1,k} \| \leq 1$ and $\| z_{j,k+1} \| \leq 1$ such that

$$e_{j+1,k} = L_{j+1,k} z_{j+1,k}, \quad e_{j,k+1} = L_{j,k+1} z_{j,k+1}. \tag{12}$$

Denoting $\zeta_{j,k} := [1, z^T_{j+1,k}, z^T_{j,k+1}, \bar{\omega}^T_{j+1,k}, \bar{\omega}^T_{j,k+1}]^T$ and taking (12) into account, Equation (11) can be rewritten as

$$e_{j+1,k+1} = \tilde{\Psi}_{j,k} \zeta_{j,k}. \tag{13}$$

Then the constraint $e^T_{j+1,k+1} R^{-1}_{j+1,k+1} e_{j+1,k+1} - 1 \leq 0$ can be expressed as

$$\zeta^T_{j,k} \big(\tilde{\Psi}^T_{j,k} R^{-1}_{j+1,k+1} \tilde{\Psi}_{j,k} - \Psi^{(0)} \big) \zeta_{j,k} \leq 0. \tag{14}$$

On the other hand, from the definition of $\bar{\omega}_{j,k}$ and constraint (2) as well as equality (12), the following constraint conditions can be derived:

$$\begin{cases} \|z_{j+1,k}\| \leq 1, & \bar{\omega}^T_{j+1,k} \mathrm{diag}\{W^{-1}_{j+1,k}, V^{-1}_{j+1,k}\} \bar{\omega}_{j+1,k} \leq 2 \\ \|z_{j,k+1}\| \leq 1, & \bar{\omega}^T_{j,k+1} \mathrm{diag}\{W^{-1}_{j,k+1}, V^{-1}_{j,k+1}\} \bar{\omega}_{j,k+1} \leq 2 \end{cases}$$

through $\zeta_{j,k}$ which can be represented as

$$\begin{cases} \zeta^T_{j,k} \Upsilon^{(1)} \zeta_{j,k} \leq 0, & \zeta^T_{j,k} \tilde{\Upsilon}^{(3)}_{j,k} \zeta_{j,k} \leq 0 \\ \zeta^T_{j,k} \Upsilon^{(2)} \zeta_{j,k} \leq 0, & \zeta^T_{j,k} \tilde{\Upsilon}^{(4)}_{j,k} \zeta_{j,k} \leq 0. \end{cases} \tag{15}$$

With the S-procedure, the inequality (14) holds if scalars $\alpha^{(r)}_{j,k} > 0$ $(r = 1, 2, 3, 4)$ exists to make the below inequality hold:

$$\tilde{\Psi}^T_{j,k} R^{-1}_{j+1,k+1} \tilde{\Psi}_{j,k} - \Upsilon_{j,k} \leq 0. \tag{16}$$

Now, we further analyze the state constraint (3). From (3), we have

$$\begin{aligned} \bar{S}_{j+1,k} \hat{x}_{j+1,k} + \bar{S}_{j+1,k} L_{j+1,k} z_{j+1,k} &= s_{j+1,k} \\ \bar{S}_{j,k+1} \hat{x}_{j,k+1} + \bar{S}_{j,k+1} L_{j,k+1} z_{j,k+1} &= s_{j,k+1} \end{aligned}$$

which can be rearranged as $\Pi^{(1)}_{j,k}\zeta_{j,k} = 0$, $\Pi^{(2)}_{j,k}\zeta_{j,k} = 0$.

Using Lemma 1, it is known that inequality (16) holds if and only if there exists $N_{j,k}$ such that

$$\tilde{\Psi}^T_{j,k} R^{-1}_{j+1,k+1} \tilde{\Psi}_{j,k} - \Upsilon_{j,k} + \mathrm{sym}\{N^T_{j,k}(\Pi^{(1)}_{j,k} + \Pi^{(2)}_{j,k})\} \leq 0. \tag{17}$$

By further utilizing the Schur complement, inequality (17) holds with the validity of inequality (10). Thus, we have prove that if $R_{j+1,k+1}$ satisfies the matrix inequality (10), with $\hat{x}_{j+1,k+1}$ being determined by (9), $\bar{x}_{j+1,k+1}$ is located in its state-estimated ellipsoid. The proof is completed.

Now, the following optimization problem is presented after establishing sufficient conditions to ensure that all possible real states would enter these ellipsoids at different shift-varying points.

Corollary 1 *Consider the 2-D DVSS (1), state equation constraint (3), the FRP and the shift-varying filter (8). If there exist matrices $K^{(1)}_{j+1,k}$, $K^{(2)}_{j,k+1}$, $N_{j,k}$ and positive scalars $\alpha^{(r)}_{j,k}$ $(r = 1, 2, 3, 4)$ for all $j, k \in \mathbb{N}$ such that*

$$\min\ \mathrm{Tr}\{R_{j+1,k+1}\} \tag{18}$$

subject to (10) is solved, then the filtering error $e_{j+1,k+1}$ always remains in the corresponding ellipsoid set that is minimized in the trace sense.

To check the achievability of the proposed SMF scheme, a SMF algorithm is developed for the FRP-based 2-D system with state equality constraints.

Algorithm 1 Recursive SMF algorithm for FRP-based 2-D system with state equality constraints

Step 1: Set $h = 1$ and the maximum step $N \in \{1, 2, \ldots\}$, the initial parameters $c_{j,k}$, $d_{j,k}$, $u_{j,k}$, $g_{j,k}$, $R_{j,0}$, $R_{0,k}$ satisfying Assumption 1 for $j, k \in \lfloor 0, N \rfloor$.
Step 2: When $j, k \in \mathbb{N}$ and $j + k = h - 1$, compute the shape matrix $R_{j+1,k+1}$, the filter gains $K^{(1)}_{j+1,k}$, $K^{(2)}_{j,k+1}$, matrix $N_{j,k}$ and the positive scalars $\alpha^{(r)}_{j,k}$ $(r = 1, 2, 3, 4)$ by solving inequality(10) and the optimization problem (18).
Step 3: Compute the matrix $L_{j+1,k+1}$ and the state estimate $\hat{x}_{j+1,k+1}$ by using (8).
Step 4: Set $h = h + 1$. If $h \leq N$, return to **Step 2**. Otherwise, go to the next step.
Step 5: Stop.

4 An Illustrative Example

This section uses a simulation example to illustrate the usefulness of the proposed filtering strategy. The system parameters of the 2-D system (1) are taken as

$$A_{j,k}^{(1)} = \begin{bmatrix} 0.5 & 0 & 0.3 & 0 \\ -0.3\sin(j+k) & -0.4 & 0 & 0.2 \\ 0 & 0.2 & 0.6 & 0 \\ -0.2 & 0 & 0 & 0.5\sin(j+k) \end{bmatrix}, \quad B_{j,k}^{(1)} = \begin{bmatrix} 0.3 \\ 0.3\sin(j+k) \\ 0 \\ 0.4 \end{bmatrix}$$

$$A_{j,k}^{(2)} = \begin{bmatrix} 0.4 & 0 & 0.3\sin(j+k) & 0.1 \\ 0 & -0.3 & 0 & 0.2 \\ 0 & 0 & 0.4\sin(j+k) & -0.2 \\ 0.2 & 0 & 0 & 0.5 \end{bmatrix}, \quad B_{j,k}^{(2)} = \begin{bmatrix} 0.23 \\ -0.16 \\ 0.2 \\ 0.32\cos(j) \end{bmatrix}$$

$$C_{j,k} = \mathrm{diag}\{0.3\sin(j+k), 0.3, 0.5\cos(k), 0.4\}$$

$$D_{j,k} = \begin{bmatrix} 0.5 & 0.4\sin(j+k) & 0 & -0.3 \end{bmatrix}^T.$$

It is assumed that the external disturbances have the following values: $\omega_{j,k} = 0.4\sin(j+k)$ and $v_{j,k} = 0.5\sin(j+k)$ with $W_{j,k} = 0.16$ and $V_{j,k} = 0.25$ for $j, k \in \lfloor 0, 30 \rfloor$. The initial states are set as $c_{j,k} = (0.12\sin(k), 0.13\cos(k), 0.1, 0.2)^T$ for $k \in \lfloor 0, 30 \rfloor$, $d_{j,k} = (0.2\cos(j), 0.2\cos(j), 0.1, 0.12)^T$ for $j \in \lfloor 0, 30 \rfloor$, $u_{j,k} = g_{j,k} = 0_{8\times 1}$ and $R_{j,0} = R_{0,k} = 0.2I$ for $j, k \in \lfloor 0, 30 \rfloor$. It can be checked that Assumption 1 holds.

In this example, we assume that the first two measurement entries are scheduled by the RRP and the last two are scheduled by the TODP, that is, $\hbar = 2$. Take the weighting coefficients $\Omega_3 = 0.3$ and $\Omega_4 = 0.5$ in the TODP. The system state constraint is supposed to be $[0\ 0\ 1\ -\sqrt{3}]x_{j,k} = 0$.

The optimization Algorithm 1 is operated by using the MATLAB software, and the corresponding parameters can be obtained as follows (for space consideration, only portion of them are listed here):

$$\alpha_{0,0}^{(1)} = 0.4434,\ \alpha_{1,2}^{(2)} = 0.1244,\ \alpha_{2,2}^{(3)} = 0.0675,\ \alpha_{2,3}^{(4)} = 0.0504,$$

$$K_{2,1}^{(1)} = \begin{bmatrix} 0.5126 & 0 & 0 & 0.7891 \\ 0.6908 & 0 & 0 & 1.1076 \\ -0.7126 & 0 & 0 & -1.2053 \\ -0.8988 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.7125 & 0 & 0 & 1.1878 \end{bmatrix}, \quad K_{2,2}^{(1)} = \begin{bmatrix} 0.5835 & 0 & 0 & 0.9056 \\ 0.4824 & 0 & 0 & 0.7922 \\ -0.4757 & 0 & 0 & -0.8095 \\ -0.3674 & 0 & 0 & -0.5882 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.3099 & 0 & 0 & 0.5157 \end{bmatrix}.$$

Based on the established recursive SMF algorithm for 2-D systems with state constraints under FRP, the filtering error $e_{j+1,k+1}$ lies in the optimal ellipsoid whose shape matrix is $R_{j+1,k+1}$.

5 Conclusion

This chapter has investigated the 2-D state-constrained system's SMF issue. During the transmission of the measurement output to the filter, the FRP has been used to decrease the phenomenon of data congestion. Combined with the method of mathematical induction, a sufficient condition has been obtained that the system state is always included in the state estimation ellipsoid at every position. A recursive algorithm for determining the optimal ellipsoids is also presented. Finally, the simulation results demonstrate the effectiveness of the proposed filtering strategy.

Acknowledgements This work was supported by the National Key Research and Development Program of China under Grant 2018AAA0100202 and the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant KYCX21_0075.

References

1. Givone DD, Roesser RP (1972) Multidimensional linear iterative circuits–general properties. IEEE Trans Comput C-21(10):1067–1073. https://doi.org/10.1109/T-C.1972.223453
2. Fornasini E, Marchesini G (1978) Doubly-indexed dynamical systems: state-space models and structural properties. Math Syst Theory 12(1):59–72. https://doi.org/10.1007/BF01776566
3. Wang F, Wang Z, Liang J, Silvestre C (2022) A recursive algorithm for secure filtering for two-dimensional state-saturated systems under network-based deception attacks. IEEE Trans Netw Sci Eng 9(2):678–688. https://doi.org/10.1109/TNSE.2021.3130297
4. Yang R, Li L, Shi P (2021) Dissipativity-based two-dimensional control and filtering for a class of switched systems. IEEE Trans Syst Man Cybern–Syst 51(5):2737–2750. https://doi.org/10.1109/TSMC.2019.2916417
5. Li M, Liang J, Wang F (2022) Robust set-membership filtering for two-dimensional systems with sensor saturation under the Round-Robin protocol. Int J Syst Sci 53(13):2773–2785. https://doi.org/10.1080/00207721.2022.2049918
6. Yang B, Qiu Q, Han Q-L, Yang F (2022) Received signal strength indicator-based indoor localization using distributed set-membership filtering. IEEE T Cybern 52(2):727–737. https://doi.org/10.1109/TCYB.2020.2983544
7. Rego BS, Scott JK, Raimondo DM, Raffo GV (2021) Set-valued state estimation of nonlinear discrete-time systems with nonlinear invariants based on constrained zonotopes. Automatica 129:109638. https://doi.org/10.1016/j.automatica.2021.109638
8. Zhu K, Wang Z, Chen Y, Wei G (2021) Neural-network-based set-membership fault estimation for 2-D systems under encoding-decoding mechanism. IEEE Trans Neural Netw Learn Syst. https://doi.org/10.1109/TNNLS.2021.3102127
9. Zhu K, Wang Z, Han Q-L, Wei G (2021) Distributed set-membership fusion filtering for nonlinear 2-D systems over sensor networks: an encoding-decoding scheme. IEEE T Cybern. https://doi.org/10.1109/TCYB.2021.3110587
10. Xu L, Li XR, Duan Z, Lan J (2013) Modeling and state estimation for dynamic systems with linear equality constraints. IEEE Trans Signal Process 61(11):2927–2939. https://doi.org/10.1109/TSP.2013.2255045
11. Ricco RA, Teixeira BOS (2022) Least-squares parameter estimation for state-space models with state equality constraints. Int J Syst Sci 53(1):1–13. https://doi.org/10.1080/00207721.2021.1936273

12. Barbosa HJC, Bernardino HS, Angelo JS (2019) An improved differential evolution algorithm for optimization including linear equality constraints. Memet Comput 11:317–329. https://doi.org/10.1007/s12293-018-0268-3
13. Ciuonzo D, Aubry A, Carotenuto V (2017) Rician MIMO channel- and jamming-aware decision fusion. IEEE Trans Signal Process 65(15):3866–3880. https://doi.org/10.1109/TSP.2017.2686375
14. Zhu K, Hu J, Liu Y, Alotaibi ND, Alsaadi FE (2021) On ℓ_2-ℓ_∞ output-feedback control scheduled by stochastic communication protocol for two-dimensional switched systems. Int J Syst Sci 52(14):2961–2976. https://doi.org/10.1080/00207721.2021.1914768
15. Ge X, Han Q-L, Zhang X-M, Ding L, Yang F (2020) Distributed event-triggered estimation over sensor networks: a survey. IEEE T Cybern 50(3):1306–1320. https://doi.org/10.1109/TCYB.2019.2917179
16. Li W, Niu Y, Cao Z (2022) Event-triggered sliding mode control for multi-agent systems subject to channel fading. Int J Syst Sci 53(6):1233–1244. https://doi.org/10.1080/00207721.2021.1995527
17. Tang Y, Zhang D, Ho DWC, Qian F (2019) Tracking control of a class of cyber-physical systems via a FlexRay communication network. IEEE T Cybern 49(4):1186–1199. https://doi.org/10.1109/TCYB.2018.2794523
18. Liu S, Wang Z, Wang L, Wei G (2022) Recursive set-membership state estimation over a FlexRay network. IEEE Trans Syst Man Cybern–Syst 52(6):3591–3601. https://doi.org/10.1109/TSMC.2021.3071390
19. Wang W, Nešić D, Postoyan R (2017) Observer design for networked control systems with FlexRay. Automatica 82:42–48. https://doi.org/10.1016/j.automatica.2017.03.038
20. Skelton RE, Iwasaki T, Grigoriadis KM (1998) A unified algebraic approach to linear control design. Taylor & Francis, Bristol, PA. https://doi.org/10.1201/9781315136523

Cloud-Based Simulation Model for Agriculture Big Data in the Kingdom of Bahrain

Mohammed Ghanim and Jaflah Alammary

Abstract The Kingdom of Bahrain has recognized big data as a power for enhancing the productivity and sustainability of agriculture as well as an essential key in developing a modern agricultural strategy that suits their climatic, water, and soil conditions. This chapter presents a simulation model for national agriculture Big Data in the Kingdom of Bahrain. The model consists of six modules (Soil, Crops, Weather, Farms, Stakeholders, and Market) with a detailed description and focus on the weather module, including tools and technologies used for the simulation, such as google cloud services, custom APIs, a Progressive Web App (PWA), and smart agriculture devices and technologies. The simulation model revealed major challenges facing the agriculture data applications in the Kingdom of Bahrain, including the absence of adequate methods and tools for data collection, inefficient storage procedures, poor data access interfaces, and media, social, and organizational limitations on data sharing, and the legal restrictions on certain automated technologies used for data collection such as the UAVs. These limitations and challenges are recommended to consider further studies from the policies and regulations side. The current study is part of a Big Data in agriculture project for the Kingdom of Bahrain, "AGRO Big Data: Toward Smart Farming in the Kingdom of Bahrain," and considers one of the few studies that addresses agriculture Big Data in the Kingdom of Bahrain.

Keywords Data engineering · Sustainable development · Smart agriculture · Data pipelines · Data warehousing · Data lake

1 Introduction

Data is an essential tool to extract insights and predictions to improve all aspects of every sector. Data is generated from different sources such as mobile phones, cars, sensors, legacy documents, structured databases, and social networks. Furthermore,

M. Ghanim (✉) · J. Alammary
University of Bahrain, Sakhir, Kingdom of Bahrain
e-mail: mghanem@uob.edu.bh

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_59

these generated data have different formats that would be only partially compatible with traditional relational database systems. This enormous amount of generated data is so-called "Big Data" and is characterized by its massive size, diversity, and high generation rate [1].

According to [2], "Big Data is the Information asset characterized by such a High Volume, Velocity, and Variety to require specific Technology and Analytical Methods for its transformation into Value." and in many contexts, it is represented by five big Vs (Volume, Variety, Velocity, Veracity, and Value) [3].

Big Data has been connected to many existing and new technologies, including data lakes, data warehouses, cloud computing, Extract, Transform, Load processes (ETL), and data pipelines. Its applications are implemented using traditional data management approaches, starting with data collection, processing, storage, analytics, and visualization. However, the unique nature of Big Data meant to handle different data sources (structured, semi-structured, and unstructured) forced some of the traditional approaches to evolve.

ETL is a traditional approach used to collect data from sources, then it runs different transformations, and cleans the data to store it in traditional repositories called data warehouses (schema-on-write). The stored data are now ready for analytics; however, this approach cannot efficiently handle semi-structured and unstructured data. Therefore, extract, load, and transform (ELT), a new approach emerged to enable the collection and storage of heterogeneous data directly on repositories is called data lakes. Data are stored in their raw format (images, text, files, etc.) [4]. Moreover, [5] suggested an extension to the approach where analysis is included in the process for Business Intelligence (BI) solutions as the following extract, load, transform, and analyze (ELTA). The data lake is a new storage concept introduced in 2010 by the CEO of Pentaho as a solution to the increased complexity and diversity of data generated [6]. Data lakes provide an unconditional repository for data, and it can custom-tailor parts of these data to form a valuable piece of information based on user requests.

The emergence of cloud computing mitigated many computation barriers, such as storage limitations, infrastructure maintenance overhead, and advanced analytical operations. Many large tech companies offer an abundance of cloud services, such as Amazon Web Services (AWS) by Amazon, Google Cloud Platform (GCP) by Google, and Azure by Microsoft Corporation. Cloud service providers offer their customers a competitive advantage on price by adopting the pay-as-you-use model, service reliability, and resilient infrastructure suitable for small, medium, and large businesses [7].

As per [8], Big Data will positively impact all food-related processes (farm to fork). Therefore, Agriculture Big Data models and applications have been of great interest to researchers. A significant number of studies discussed applications for crops [9], weed control [10], precision agriculture [11], etc. Most of the models presented were focusing on analytics, machine learning, and the internet of things; however, few studies focused on agriculture Big Data engineering operations (data collection, processing, and storage), and up to the authors' knowledge, agriculture

Big Data engineering implementation on GCP has not been discussed by previous researches.

The Kingdom of Bahrain severely lacks agriculture data. Data are either scattered, outdated, or mostly not available. A sustainable agriculture system relies heavily on data to create knowledge that will eventually lead to effective decisions. Big data analysis has proven effective in many fields, including the agriculture sector. However, building and maintaining a Big Data infrastructure is costly, time-consuming, and depletes resources.

The study is part of an experimental project on agriculture Big Data for the Kingdom of Bahrain that will focus on the data engineering process starting from data collection and ending by delivering analytical-ready data. The study will concentrate on one of the modules of the ecosystem which is weather data and its implementation process by (1) defining data sources, (2) illustrating tools used for data ingestion, (3) defining scalable data pipelines that efficiently accommodate both streaming and patch data, (4) finally, demonstrating the architecture of the storage schema. The Weather module was selected mainly for its fundamental impact on the farming practices for crops. Data pipelines are used to model the data flow and technologies used to deliver data from source to destination.

The following sections will present, in order, a literature review of related work, the tools and the approach used for implementation, a discussion of the implementation through briefly explaining the entire simulation, and then a detailed description of the weather module, and finally, a conclusion and future work.

2 Research Background

Data engineering studies related to agriculture Big Data focusing on data ingestion, transformation, and storage are not sufficiently presented. Available studies either demonstrate the complete ETL/ELT processes or focus on only one of the processes, such as the storage or transformation of data.

Data transformation is where data is cleaned and converted to a suitable format for analysis. [12] Used Couchbase (NoSQL tool) to transform unstructured agriculture Big Data into semi-structured or even structured data. The model suggests the collection of agriculture data from various unstructured sources such as TEXT and XML then applying MapReduce (Distributed Data Processing Algorithm) algorithm on Couchbase to convert it into structured data.

Storage technologies improved and varied significantly during the past few years—the advancement in data repositories derived mainly from the increased complexity and heterogeneity of data generated. Storage operations are now far from being a simple task; it is cloud-based with auto-scaling and the availability of enormous capabilities at the user's fingertips. Establishing an efficient and reliable data warehouse is essential to perform business intelligence analysis on data; however, a data lake must be used as a preprocessing step to handle raw data [13]. Vuong et al. [14] designed a data integration module for agriculture Big Data by combining

Apache Hive (as a data warehouse and running OLAPs), MongoDB (for document storage), and Apache Cassandra as a data lake for raw data.

Full implementation of data engineering processes in agriculture Big Data was found on [15]. The study is based on a platform built and deployed on five machines connected through TCP/IP. Apache Flume and Sqoop were used to collect real-time data and import data from CSV and Excel files. For large files, MapReduce algorithms were adopted to collect them in parallel, and then the collected data are stored either on HDFS for archival or Hbase and Hive for random access and real-time analytics.

On the other hand, agriculture Big Data analytics, which involves extracting the value of agriculture data sets using machine learning, neural networks, and data mining, has been studied heavily. The concept behind Big Data paved the way for many agriculture-related applications and models, many of which are related to analytics to extract value for better decisions.

Neural Networks (NN), Support Vector Machine (SVM), and Graphical Models are, according to [10], the most used machine learning algorithms for agriculture data analysis. Moreover, the study suggested a conceptual model using both SVM and NN algorithms to identify crops' weeds and potential diseases by labeling and classifying field images collected using Unmanned Aerial Vehicles (UAVs) in the Netherlands. GreenLink is a mobile app that will analyze data collected using six Wireless Sensor Networks (WSNs) in a small test farm. The sensors will collect data about water, weather, energy consumption, and crops, then store the data on the Azure cloud platform for analysis using deep neural networks and regression trees algorithms [11]. Data mining algorithms such as PAM, CLARA, and DBSCAN can provide wheat farmers in India with valuable insights to minimize inputs (resources) and maximize production [9].

3 Methodology

To develop agriculture Big Data applications for Bahrain, a simulation model is adopted to imitate the processes of data ingestion, transformation, and storage. The simulation model consists of six modules (Weather, Soil, Crops, Market, Stakeholders, and Farms). The modules were suggested based on the agriculture ecosystem in Bahrain and related literature. It includes the development of a Progressive Web App (PWA) to simplify communication with different agriculture stakeholders as well as the collection of data. The PWA is hosted on google firebase to streamline communication with other google cloud services. Various tools and technologies have been utilized (see Table 1); however, the main workflow is implemented within the GCP.

Google Cloud Platform is selected for implementation as a comprehensive tool to run an experimental simulation of an agriculture Big Data ecosystem on the cloud and not based on comparison with other cloud service providers. It provides many services that support complete data engineering and data analysis operations. Moreover, it

Table 1 List of tools and technologies adopted for the simulation model

	Tool/Technology	Quantity	Description
1	Custom built PWA	1	The app enables data collection for demographic data of stakeholders, farms' data, adding new data sources, and it provides an Agriculture Enquiry Service (AES) functionality. On the other hand, it will provide easy access to agriculture data, and to enable collection of legacy data such as excel files and as a hub to import individual projects' data
2	Pycno soil sensor	1	A field-based sensor to collect soil and weather data, it can be installed in open and protective fields to collect and transmit data using cellular connectivity
3	Remote sensing account	1	An account created on Earth Observation System (EOS) crop monitoring service to remotely collect data on farm's weather, soil moisture, and NDVI index images. The satellite images are captured using SENTINEL-2 and Landsat 8 satellites
4	Multispectral camera	1	Due to policies and regulations, the camera was used manually (by hand) to capture images of crops to mimic the drone-based cameras
5	Google cloud platform (GCP)	1	A google account was created to take advantage of cloud services such as (BigQuery, cloud storage, cloud functions)
6	Visual studio code	1	The main development environment for the simulation model
7	Google firebase	1	Used to host the PWA and the use of the NoSQL capabilities of Firestore
8	Front-end development using Ionic framework	1	Ionic framework is well-known development environment for developing hybrid mobile applications and PWAs
9	Backend development using python and flask framework	1	Python is an easy high-level programming language with strong support for data science. Flask is a python web framework adopted by google cloud services

offers highly scalable and efficient storage tools that suit data lakes and warehouse operations [16].

The weather module implementation will be discussed in detail to present the different pipelines used to complete the process of ETL/ELT on the data and the used data schema. Six GCP services were mainly adopted for the weather module:

- Google BigQuery: Cost-effective, serverless, multi-cloud enterprise data warehouse for structured Big Data storage and analytics.
- Cloud storage (Data Lake): In GCP, cloud storage is a scalable and real-time storage repository that can handle both semi-structured and unstructured data with easy access operations.
- Cloud SQL (MySQL): Fully managed relational database service for MySQL for storing structured data.

- Serverless Cloud Functions: Cloud-based functions to connect or extend cloud services.
- Google Pub/Sub: Messaging-oriented middleware for service integration or as a queue to parallelize tasks.
- Cloud Scheduler: Fully managed enterprise-grade cron job scheduler to schedule pub/sub messaging services.

The agriculture sector in Bahrain comprises many sources for agriculture data; however, due to organizational, technical, and social limitations, this data is outdated, not publicly available, difficult to access, difficult to share, and in some situations, does not exist. According to the personal interviews and field studies, the following are considered the primary agriculture data sources in Bahrain:

1. Ministry of Municipalities Affairs and Agriculture-Agriculture and Marine Resources is the main governmental entity responsible for the agriculture sector in Bahrain. It owns many data, including laws and legalizations, market data, soil and water data, and crop production and quality. Market data are collected and stored using a particular platform; however, it is not available publicly. Soil and water data used to be collected manually by official staff and primarily stored on excel or even word files on local machines, and these data are becoming outdated due to a shortage of staff. On the other hand, the ministry installed two weather stations in the main agriculture areas (Hoorat A'ali and Budaiya); however, the data are neither accessible nor shareable with different agriculture stakeholders.
2. The National Initiative for Agriculture Development (NIAD) is a non-governmental organization that supports the agriculture sector, especially farmers. They launched in the second quarter of 2022 a website (https://www.agro.bh/) to share agriculture data they collected, such as available farms, vendors, and farmers. It is considered the first initiative to collect and share agriculture data in Bahrain and will be used as a web scrapping source on the simulation model.
3. The Bahrain data portal (https://data.gov.bh/) is an official portal for open data about many sectors in Bahrain. It contains legacy data about Bahrain weather; however, for technical limitations, the data cannot be exported in formats such as (CSV or JSON). Moreover, the portal does not provide API documentation so that third parties can use the data.
4. The National Space Science Agency (NSSA) was established in 2014 to support the achievements of the sustainable development goals in Bahrain by adopting space sciences and satellite technologies. They run agriculture data projects; however, data is neither accessible publicly nor anonymously for research purposes.
5. Individual data owners such as farmers, farm owners, and even agronomists or agriculture scientists have run their projects where they collect data either

manually or using technology; however, the data they collect is neither stored efficiently nor available for sharing because of technical and social limitations.

According to the above, data sources identification, demand data, and agriculture data beneficiaries were simulated based on the information collected through:

- Eleven personal interviews using semi-structured questions during the period April 2022–September 2022—The interviews were conducted with selected interviewees based on the agriculture ecosystem in Bahrain to cover all stakeholders (farmers, governmental organizations, farm/business owners, and non-governmental organizations, and agronomists). The interview questions focused on the data collection method and tools, storage of the data, is the data shared, methods to access the data, and difficulties in obtaining the required data.
- Five personal field visits, two of which were an IoT device, were used as part of the data sources simulation between November 2021 and January 2022.
- Related literature.

Finally, a custom Big Data management model was applied and divided into three layers as follows (see Fig. 1):

- Bottom layer: a Meta catalog containing technical metadata about the data definition and structure such as (tables, fields, and their description, and relation between tables).
- Middle layer: business metadata such as data sources and description of data, supportive tools, data movements.
- Top layer: this layer is further divided into two structures, the first one will contain sources-specific data generated, and the second is an aggregation of all sources based on data category. For example, weather data can be generated using soil sensors, remote sensing, and weather stations so that every source will have

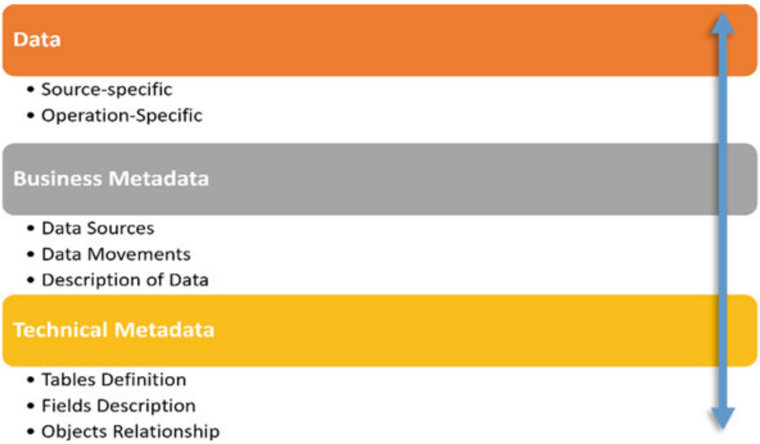


Fig. 1 The conceptual agriculture big data management model

its structured/unstructured storage mechanism. Then, all weather data will be transformed, aggregated, and stored on operation-specific tables.

4 Design and Implementation

Due to organizational and funding limitations, the simulation model is built on both real and mimicked data sources. The following sub-sections will demonstrate a general overview of the processes (ingestion, storage (raw and processed), transformation) and, finally, a detailed demonstration of the weather module for all processes.

4.1 Data Ingestion

Data ingestion is the process of collecting data from numerous sources, such as data streams from events, logs and IoT devices, historical data stores, and data from transactional applications. However, according to [17], a preliminary stage is adopted to identify data providers and users to improve source quality and increase data privacy. Table 2 represents a matrix of agriculture stakeholders in Bahrain (data owners and data users) and the categories of data they can record/access for the agriculture sector. The matrix is built based on the agriculture sector ecosystem in Bahrain. Google cloud platform offers many tools to orchestrate the ingestion process of data, including (cloud scheduler, pub/sub, and cloud functions/run).

Two metadata tables were created, and the stakeholder's table will inherit access level from them, and data were predefined based on the matrix as follows:

Data Sources Identification and Collection

The simulation model aims to collect different agriculture data semi-autonomously and then build scalable pipelines to deliver the data to the cloud for storage, processing, and preparation for analytical operations. Data sources were identified based on personal interviews with different agriculture stakeholders, field visits, and related literature. Table 3 identifies data sources used by the simulation model; most data are collected through various sources. However, a stakeholder account must be created through the AgroBahrain PWA forms to initiate the collection and any other related functionality. To automate the real-time data collection, an IoT soil sensor (Pycno) was used, and a subscription to a remote sensing application based on satellite imaging (EOS) that provides crop monitoring services (see Fig. 2).

Table 2 Agriculture data owners'/users' matrix for the Kingdom of Bahrain [18]

		Data owners					
		Farm owner	Farmer	Gov. official	Non-Gov. official	Merchant	Researchers, agronomists
Data users	Farm owner	Personal, farms' data, crops' data, sensors' generated, market data, experts' data	Personal, experts' data	Personal, policies and regulations, sensors' generated, market data, experts' data	Personal	Market data	Experts' data
	Farmer	Farms' data, crops' data, sensors' generated, experts' data	Personal, experts' data	Personal, policies and regulations, sensors' generated, experts' data	Personal		Experts' data
	Gov. official	Personal, farms' data, crops' data, sensors' generated, market data, experts' data	Personal, experts' data	Personal, policies and regulations, sensors' generated, market data, experts' data	Personal	Personal, market data	Personal, experts' data
	Non-Gov. official	Farms' data, crops' data, market data, experts' data	Personal, experts' data	Policies and regulations, market data, experts' data	Personal	Personal, market data	Personal, experts' data
	Merchant	Crops' data, market data		Policies and regulations, market data		Personal, market data	

(continued)

Table 2 (continued)

		Data owners					
		Farm owner	Farmer	Gov. official	Non-Gov. official	Merchant	Researchers, agronomists
	Researchers (Agronomists)	Farms' data, crops' data, sensors' generated, experts' data	Experts' data	Policies and regulations, sensors' generated, experts' data			Personal, experts' data

Table 3 Data sources used for the simulation model

Data source	Data to collect	Collection mechanism/interface
AgroBahrain PWA	Demographic, farms, crops, fertilizers, etc.	Forms
Pycno soil sensor	Soil temp, Soil moisture @ 25 cm depth, Air temp, Air humidity, Solar radiation, Rainfall intensity	Cellular connectivity. Data are collected every 35 minutes and sent twice a day to the server. The API will capture the data every (hours)
Remote sensing platform (EOS, the crop monitoring App)	Weather data (temp, humidity, perception), soil moisture, vegetation indices (e.g., NDVI), images of fields of interest	SENTINEL-2 and Landsat 8 satellites, collect historical data, images are captured every 5 days, weather forecast for today and the next two days. Specific APIs are required per data category for collection
Web crawling/ scrapping	Farms, market, stakeholders, etc.	Custom python scripts. URLs such as (agro.bh)
National weather stations	Weather data	APIs, Manual upload
Other sources	Agriculture-related (real-time and legacy) data	Custom APIs, Manual upload through the PWA forms

Due to organizational and technical limitations, the following sources were simulated as follows:

- National weather stations data collection was considered as the APIs used on the IoT device or through manual upload using CSV files using the AgroBahrain platform.
- Other sources may include legacy data from governmental and non-governmental sources in a CSV format. Despite the availability of legacy metrological data on the (Bahrain Data Portal), it was not downloadable and difficult to access. Therefore, two legacy weather CSV files (1901–2021) were used from the World Bank repository (https://climateknowledgeportal.worldbank.org/country/

```
"S1T": 66.2,
"S2T": 61.81,
"HUM": 96.34,
"TEMP": 15.18,
"LW1": 3.28,
"LX1": 5896,
"RAINH": 0,
"name": "pycno 1",
"LFREQ": 868,
"NS": 320,
"BAT": 4.15,
"CME": 0,
"GR": 1,
"GRF": 0,
"VE": 4.73,
"MDM": 68,
"TXT": "No GPS",
"loc": {
   "type": "Point",
   "zoom": 14,
   "valid": 1,
   "updatedAt": "2021-12-25T06:24:46.231Z",
   "satellites": 6,
   "coordinates": [
      "50.532875000",
      "26.171695000"
   ]
},
"epoch": 1641878076662,
"polledAt": "2022-01-11T05:14:36.662
```

```
[
  {
    "2022-10-30T00:00:00+00:00": {
      "t_2m__min": 298.2,
      "t_2m__max": 300.6,
      "asob_s__sum": 0,
      "tot_prec__sum": 0,
      "h_snow__sum": 0,
      "relhum_2m__avg": 85.7,
      "u_10m__avg": -0.9,
      "v_10m__avg": 0.8,
      "w_speed__avg": 1.2,
      "w_speed__min": 1,
      "w_speed__max": 1.5,
      "clct__avg": 1.3,
      "td_2m__avg": 296.6,
      "t_so__min": 298.6,
      "t_so__max": 300.5,
      "ps__avg": 101196,
      "w_dir__avg": 312.2,
      "ww": 0,
      "ww_human": "clear sky"
    }
  },
  {
    "2022-10-30T01:00:00+00:00": {
      "t_2m__min": 297.7,
      "t_2m__max": 300.4,
      "asob_s__sum": 0,
```


Fig. 2 Sample of data collected. On the left sample JSON data from Pycno soil sensor, and on the right sample JSON of EOS satellite data

bahrain/climate-data-historical). The first file presents the average monthly temperature in Bahrain, and the second shows the average monthly precipitation.

Collection of data was implemented using scheduled python scripts triggered through pub/sub events and run using google cloud functions. The collection methods are as follows:

- The PWA forms are used to collect data about stakeholders, farms, crops, loading data sources such as legacy data, and registering automated data sources such as field's sensors and weather stations.
- Web crawling/scrapping of related URLs such as national newspapers and "agro.bh" web site.

4.2 Data Storage (Raw and Processed)

Data is either stored in raw format or for processing (analytics, querying, visualizations). Raw data are collected from their original sources (sensors, satellite images, legacy files, etc.) and will be stored on Cloud Storage (Data Lake) for the archival and preprocessing stage.

Metadata, transactions data, and other data collected through AgroBahrain PWA forms will be stored on Cloud SQL (MySQL) except for the agriculture enquiries service (AES), the data will be stored on google Firestore (NoSQL document database). Cleaned transformed data will be stored on google BigQuery (the data warehouse) for fast and reliable access to analytics or visualizations. Data for visualizations are aggregated from all sources and are anonymous; however, data for analytics are aggregated for each source and identified by source ID. The adopted storage schema is a multidimensional model called fact constellation schema. It is flexible for complex structures like agriculture data; it allows multiple joins and multiple facts and dimension tables.

4.3 Data Processing (Transformation)

Transformation is an iterative process that will be carried out by running cloud functions and custom SQL commands, and it will be divided into:

- Extracting module-related data. Some sources will collect data related to two or more modules (e.g., weather, soil, and crops) so these data must be segregated.
- Cleaning segregated data by checking for duplicates, removing extra spaces and null values, and checking for data range and extreme values.
- Unifying data types, measurement units, and fields' names.
- Aggregation of data to create visualization and analytical-ready data for stakeholders.

4.4 Implementation of the Weather Module (ELT/ETL Process)

Adding a new stakeholder (register new user): The new stakeholder must provide basic personal data through a form, and then the data will be stored in a structured database. Next, the data is inserted into a structured database on google cloud SQL (MySQL). Once registered successfully, a stakeholder can start adding data sources.

Adding new data sources (see Figs. 3 and 4): As for the simulation model, weather data will be generated through (pycno farm's sensor, satellite remote sensing, and legacy data). New data sources are either added directly, such as legacy data or related to a specific geolocation, such as sensors on a farm that cultivate certain crops. Every insert for a data source will affect three entities. The first entity will record the source data fields, data types, and measuring units (Km, L, Celsius, etc.). The second entity will record source APIs for automated sources (sensors, remote sensing, weather stations); however, this entity will not be updated for manually added data, such as legacy data. The third entity will create a google scheduler record that initiates an event using a messaging service such as (Pub/Sub), which will trigger a google cloud function to ingest the data. A stakeholder can register two types of data sources:

1. Automated data sources such as farm's related sensors and weather stations.
2. Manual data sources such as legacy data or real-time manually added data.

Extraction and storage of raw data (schema-on-read) (see Fig. 5): All raw weather data simulated are either structured (CSV files) or semi-structured (JSON) and will be stored and archived using the cloud storage buckets. All the files will be named based on the (source_id), and a new field will be added to every file (created_at) as a time stamp for every data record. Data were extracted and stored either directly or indirectly as follows:

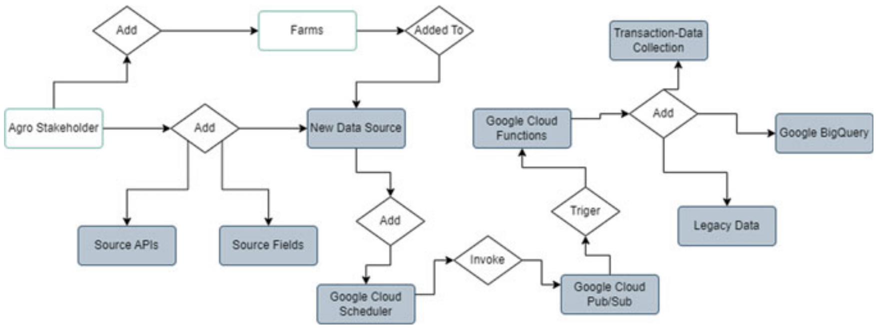


Fig. 3 Logical data flow diagram for the weather module

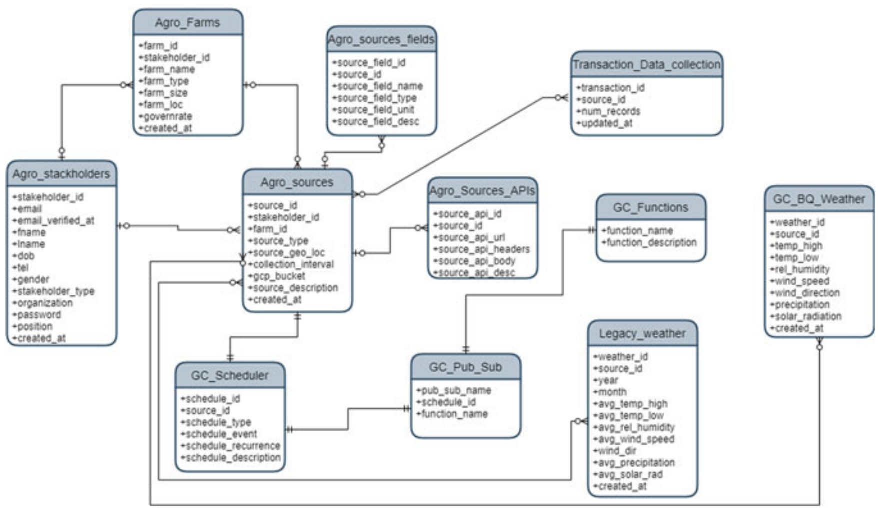


Fig. 4 The weather module data scheme

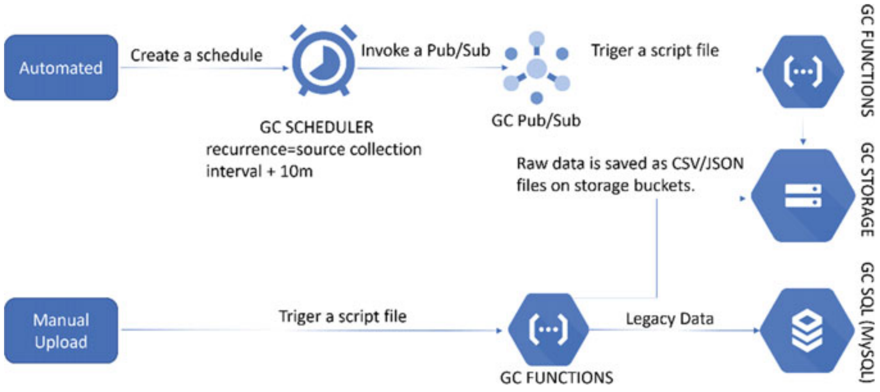


Fig. 5 Extraction and storage of raw data pipeline

- Manually added data, such as legacy weather data is created and stored directly through the PWA form for adding new data sources. It will accept either CSV or Excel files only. On the backend, a python script will be executed by cloud functions (see Table 4) to extract the uploaded file and store it under the cloud storage bucket (e.g., legacy_weather).
- Automated sources data such as farm sensors and weather stations are extracted and stored based on scheduled pub/sub events that trigger a python script to

Table 4 Cloud functions used for the weather module

Cloud Function	Description
Add_ Stakeholder	a. Insert a new stakeholder record in cloud SQL table (Agro_Stakeholders)
Add_Data_ Source	a. Insert a new data source by updating the following tables (Agro_sources, Agro_sources_APIs, Agro_sources_Fields, GC_Scheduler, GC_Pub_Sub) b. Create CSV/JSON files on cloud storage buckets c. Create new cloud schedule and new Pub/Sub d. Insert CSV/Excel files directly to cloud storage (only legacy data) e. Transform legacy data and insert into cloud SQL (Legacy_Weather) table
Raw_Data_ Extract	a. The function is either called manually by PWA form (case of manual data upload) or through a Pub/Sub b. Update the CSV/JSON files c. Update the table (Transactions_data_collection)
BigQuery_ weather	a. The function is triggered every day to collect the recently added records on all CSV/JSON b. Extract fields with related weather tags (temp, temperature, temp_high, hum, humidity, etc.) c. Cleaning and transforming the extracted data d. Insert the data into BigQuery Tables (GC_BQ_Weather_Agreggated and GC_BQ_Weather_Detail)

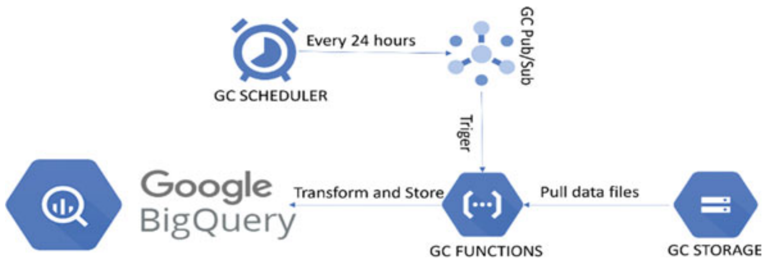


Fig. 6 Transformation and storage (analytics and visualizations) of data pipeline

extract and insert the data into CSV/JSON files (based on the source format) to be stored in a cloud storage bucket named according to the source type (farm_sensor, weather_stations, etc.) to increase the search efficiency. The CSV/JSON file is created once the data source is added and will be updated every time the script runs.

A Meta table (Transaction_Data_Collection) is used to track data extraction and storage transactions. Legacy data will be extracted and stored once and is registered only on the data sources table.

Transformation and Storage of Processed Data (schema-on-write) (Fig. 6): Every 24 hours, a cloud function (see Table 4) will be triggered to extract updated data from the cloud storage, transform the data, and then store the transformed data on google BigQuery tables for fast analytics and visualizations. For legacy data, they are transformed, stored in a Cloud SQL table, and accessed through BigQuery federated queries (data is handled efficiently without the need to store it physically in a BigQuery table).

5 Conclusion

Big Data has become a hot topic in the last decade, with many countries looking to acquire a competitive edge in this new industry. The Kingdom of Bahrain government has recognized the power of Big Data to drive economic growth, enhance decision-making, and create competitive advantages in both the public and private sectors. In agriculture, Big Data will accurately detect wind direction, temperature, relative humidity, solar and evaporation, and plant transpiration. In general, Big Data is significant for improving and sustaining the agriculture sector in the Kingdom of Bahrain. However, Bahrain lacks agriculture data with low adoption of ICT in agriculture. Therefore, implementing a Big Data model for such a situation is complex and challenging.

The current study is part of a Big Data in agriculture project for Bahrain, "AGRO Big Data: Toward Smart farming in the Kingdom of Bahrain." In the study, a simulation model with six modules for agriculture Big Data (Weather, Soil, Crops, Market, Stakeholders, and Farms) was demonstrated briefly while focusing on the implementation details of the weather module. The simulation was implemented to imitate the data engineering processes for collection, transformation, and storage. The simulation model revealed many challenges hamper the agriculture Big Data development in Bahrain, including the absence of adequate methods and tools for data collection, inefficient storage procedures, poor data access interfaces, and media, social, and organizational limitations on data sharing, and the legal restrictions on certain automated technologies used for data collection such as the UAVs. Moreover, there is a lack of ICT knowledge and adoption among the agriculture stakeholders, which presents a significant barrier toward implementing agriculture Big Data solutions.

The simulation model was designed for Bahrain; however, it was built on the google cloud platform (GCP) to be flexible enough to suit other countries. GCP offers accessible serverless cloud services that can scale automatically, and it accommodates all the simulation model processes without needing external services.

For future work, the simulation model will be expanded to enable multi-access levels for agriculture stakeholders to the collected data, where everyone can benefit and reuse the available data while maintaining efficient privacy and security levels. Furthermore, collected data will be transformed into valuable application-specific data sets for decision-making and research purposes.

References

1. Hashem IA, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU (2015) The rise of "big data" on cloud computing: review and open research issues. Inf Syst 47:98–115
2. Mauro AD, Greco M, Grimaldi M (2016) A formal definition of Big Data based on its essential features. Libr Rev 65(3):122–135
3. Sarker MN, Islam MS, Murmu H, Rozario E (2020) Role of big data on digital farming. Int J Sci Technol Res 1–11
4. Zouari F, Kapachi NG (2021) Data management in the data lake: a systematic mapping. In: 25th International database engineering and applications symposium. ACM, Montreal, pp 280–284
5. Marín-Ortegaa PM, Dmitriyev V (2014) ELTA: new approach in designing business intelligence solutions in era of big data. Proc Technol 16:667–674
6. Khine PP, Wang ZS (2018) Data lake: a new ideology in big data era. In: 4th Annual international conference on wireless communication and sensor network (WCSN 2017). EDP Sciences, Wuhan, pp 395–405
7. Tabish Mufti PM (2021) A review on amazon web service (AWS), Microsoft azure & google cloud platform (GCP) services. In: 2nd international conference on ICT for digital, smart, and sustainable development (ICIDSSD 2020). EAI, New Delhi
8. Magnin C (2016) How big data will revolutionize the global food chain. Retrieved from McKinsey & Company: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/how-big-data-will-revolutionize-the-global-food-chain
9. Majumdar J, Naraseeyappa S (2017) Analysis of agriculture data using data mining techniques: application of big data. J Big Data 4

10. van Evert FK, Fountas S, Jakovetic D, Crnojevic V, Travlos I, Kempenaar C (2017) Big data for weed control and crop protection. Weed Res 57:218–233
11. Yemeserach Mekonnen SN (2020) Review—machine learning techniques in wireless sensor network based precision agriculture. J Electrochem Soc 167(3)
12. Sambrekar K, Rajpurohit VS, Joshi J (2018) A proposed technique for conversion of unstructured agro-data to semi-structured or structured data. In: Fourth international conference on computing communication control and automation. IEEE, Pune, pp 1–5
13. Neves RA, Crunivel PE (2020) Model for semantic base structuring of digital data to support agricultural management. In: 14th International conference on semantic computing. IEEE, California, pp 337–340
14. Ngo VM, Le-Khac N-A, Kechadi M-T (2019) Designing and implementing data warehouse for agricultural big data. In: Keke Chen SSJ, Big data—BigData 2019. Springer, California, pp 1–17
15. Nguyen V-Q, Nguyen SN, Kim K (2017) Design of a platform for collecting and analyzing agricultural big data. J Digit Contents Soc 18(1):149–158
16. Bisong E (2019) An overview of google cloud platform services. In: Bisong E, Building machine learning and deep learning models on google cloud platform. Apress, OTTAWA, p 709
17. Paolo Ceravolo AAMU (2018) Big data semantics. J Data Semant 7:65–85
18. Ghanim ME, Alammary J (2021) AgroBahrain: a conceptual framework for agriculture big data for Bahrain. In: International conference on electronic information technology and smart agriculture. IEEE, Huaihua, pp 545–550

User Interface Design and Evaluation of the INPACT Telerehabilitation Platform

Leonor Portugal da Fonseca, Renato Santos, Paula Amorim, and Paula Alexandra Silva

Abstract With longevity, sedentarism and the increasing strain on health services, telerehabilitation has gained increasing importance. This paper describes the user interface (UI) design, development and evaluation of a telerehabilitation platform. Following a human-centred design approach, two UIs were developed: one for persons undergoing rehabilitation (PUR) and another for rehabilitation professionals (RP). The usability of these UIs has been assessed by two groups of five users, one who tested the PUR UI prototype and another who tested the RP UI prototype. After completing a set of tasks, on which number of errors, task duration, completion, and utility have been recorded, participants answered the System Usability Scale and the Computer System Usability Questionnaire. To get the likelihood of participants recommending the system participants filled out the Net Promoter Score. This methodology has shown to be useful to fine-tune the initial user requirements of the system and evaluate the UIs developed and has shown the importance of involving the several actors in the design of the platform.

Keywords Usability · Telerehabilitation · Design · Evaluation

1 Introduction

Physical rehabilitation has been at the forefront of controlling and improving a diversity of health conditions, from musculoskeletal to neurological, cardiorespiratory and others. Among these, musculoskeletal conditions are the most prevalent, affecting about 1.7 billion people worldwide [1]. With the progressive ageing of Western soci-

L. P. da Fonseca (✉) · R. Santos · P. A. Silva
Department of Informatics Engineering, University of Coimbra, Centre for Informatics and Systems of the University of Coimbra, Coimbra, Portugal
e-mail: uc2019183980@student.uc.pt

P. Amorim
Faculty of Health Sciences, University of Beira Interior, Covilhã, Portugal

Rehabilitation Medicine Center of Central Region, Monte Redondo, Portugal

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_60

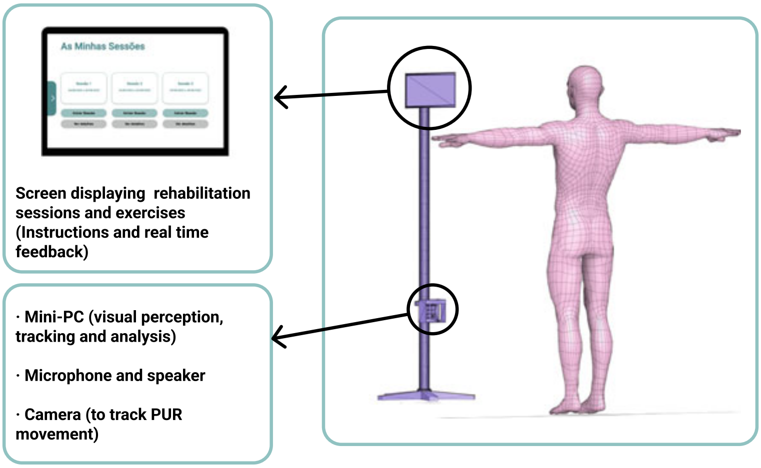


Fig. 1 Telerehabilitation platform components

eties and the increase in sedentary lifestyles, musculoskeletal conditions are expected to increase [2]. Physical rehabilitation may help treat and alleviate symptoms, but, in order for it to be effective, it requires a continuous process that is very demanding both in terms of human resources and financial costs, which will become unsustainable for the health system as demand continues to increase [3]. Telerehabilitation provides a way to mitigate this problem, as rehabilitation sessions can take place outside health units, such as hospitals and clinics, while optimising processes and avoiding unnecessary travels [4]. Further, telerehabilitation could allow for the reduction of waiting lists and the delivery of physiotherapy care to people in remote and underprivileged areas [5].

The INPACT project aims to create a low-cost remote rehabilitation platform that, through the use of a camera, allows a holistic visual perception of the user's body movement without the use of markers (Fig. 1). Two main users interact with the system: the person undergoing rehabilitation (PUR), at home, and the rehabilitation professional (RP), at the health unit they usually work. The PUR user interface displays a series of sessions and exercises tailored to the person by the RP on his/her dedicated user interface. As the PUR performs the exercises prescribed by the RP, the camera captures the PUR movement-related data. This data is then analysed using machine learning techniques that enable the provision of real-time feedback to the PUR. The data collected during the sessions is stored in the cloud, allowing the RP to monitor and analyse the performance of the PUR and to adjust the process as needed.

The INPACT system followed a human-centred design (HCD) process [6]. An iterative and incremental process was followed for the development of the prototypes. Having achieved a stable low-fidelity prototype, the team started implementing a web application that uses React.JS, where the communication between the frontend and the backend is supported by a REST API that responds to the user request, resorting to cloud services (Fig. 2).

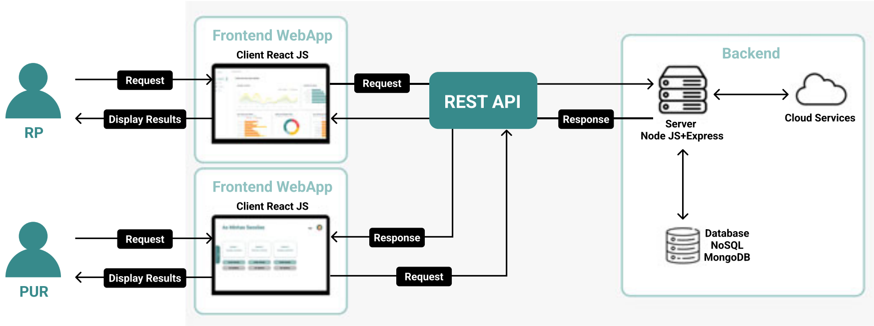


Fig. 2 General system architecture

2 Evaluation of the User Interface Prototypes

2.1 Goals, Methods and Procedures

Part of a larger effort, this paper describes the usability evaluation of the first high-fidelity prototypes developed for the PUR and the RP which had been deployed on Vercel.[1] The goal of this evaluation was to assess the usability of the user interfaces of the PUR and RP and to get a sense of the participants' views on the usefulness of the functionalities included in the prototypes as well as the likelihood of them recommending the system to others. Usability was assessed i) by asking participants to go through a set of tasks and recording its duration, number of errors, and task completion and ii) through a post-task self-report usability questionnaire. Table 1 lists the tasks that guided the usability test. After completing each task, participants were prompted to provide comments or suggestions. To gauge the perceived usefulness of each task, participants were asked to rate the usefulness of the task by using an 8-point scale between 0 (not at all useful) and 7 (very useful).

Once all usability tasks were completed, post-sessions ratings were collected to assess the overall perceived usability. We used the System Usability Scale (SUS) [7] to assess the PUR prototype and the Computer System Usability Questionnaire (CSUQ) [8] to assess the RP prototype. SUS is one of the most widely used tools for assessing usability [9] and would have been adequate for both prototypes, but CSUQ includes questions that directly analyse matters such as productivity, which we considered particularly relevant for the RP. For this reason, we applied the CSUQ with the participants assessing the RP prototype. Both SUS and CSUQ provide reliable measures to assess the user impressions of the system [10]. Finally, we used the Net Promoter Score (NPS), a popular metric to get the extent to which a person would recommend a system [9].

[1] https://vercel.com/.

Table 1 List of tasks the participants were invited to complete

PUR prototype test tasks (T)	RP prototype test tasks (T)
T1. Check the details of the rehabilitation session scheduled for today	T1. Identify the PUR who has been experiencing the highest level of fatigue
T2. Change the settings to choose preferred background	T2. Check the performance of a PUR
T3. Start rehabilitation session and check number of exercises remaining in session and number of exercises repetitions	T3. Check the level of fatigue of a PUR when performing a specific exercise
T4. Assess overall session and fatigue and pain post exercise	T4. Create a general session plan
T5. Leave a message to the RP	T5. Change repetition parameters of exercise
T6. Quit the session halfway through	T6. Assign existing session plan to PUR

The evaluation was carried out by two members of the team, one as an observer, mostly for note-taking, and another as a facilitator, who guided participants through the test. The facilitator started by providing a brief description of the system, explaining that it had the PUR and the RP sides. Then the facilitator clarified that the test was going to focus only on the PUR or the RP user interface, depending on the profile. After this, informed consent was gathered and the facilitator collected socio-demographic information about the participants. In the case of the PUR participants, the facilitator also asked about previous history of undergoing rehabilitation. To the RP, the facilitator asked about their number of years of experience as a rehabilitation professional. Afterwards, the usability test started, where the facilitator first invited the participants to freely explore and browse the user interface for about 2–3 minutes. The facilitator then asked the participants to consider a scenario and complete each of the usability tasks (Table 1), encouraging them to interact with the user interface prototype as naturally as possible. Once the last usability task was completed, the participants were asked to fill out the post-session questionnaires and the NPS. The test finished with the facilitator thanking the participants.

2.2 Participants

The evaluation involved five participants who tested the PUR user interface and five who evaluated the RP side of the system. The participants who tested the PUR user interface were 25–65 years of age and had all previously received conventional physical rehabilitation sessions. They also reported on the daily use of a tablet or smartphone and the regular use of these devices to email, make video calls and using tools like MS Office. The participants who tested the RP user interface prototypes were all physical rehabilitation professionals with 2–5 years of experience in the job and aged 18–30. These participants reported using computers to check email, news,

Table 2 Results of the tests on the PUR user interface prototype

Participant	Task 1			Task 2			Task 3			Task 4			Task 5			Task 6		
	D	E	U	D	E	U	D	E	U	D	E	U	D	E	U	D	E	U
PURTP1	41s	2	5	26s	3	6	46s	3	6	11s	1	5	21s	0	6	13s	0	6
PURTP2	31s	2	6	24s	1	5	54s	1	7	25s	2	7	24s	1	7	19s	0	7
PURTP3	54s	2	5	58s	3	7	30s	0	7	10s	0	6	31s	2	7	38s	0	7
PURTP4	18s	1	7	10s	1	6	34s	1	7	9s	0	7	7s	0	7	16s	1	6
PURTP5	25s	1	6	16s	1	6	38s	0	7	12s	1	5	13s	1	5	17s	0	6
Mean	33.8s	1.6	5.8	26.8s	1.8	6	40.4s	1	6.8	13.4s	0.8	6	19.2s	0.8	6.4	20.6	0.2	6.4

Legend PURTP—person undergoing rehabilitation test participant, D—task duration, E—errors, U—usefulness

weather and sometimes to make video calls and work on tools like MS Office. None of the participants involved in the study had previously been exposed to the INPACT prototypes.

3 Results

3.1 Usability Tasks and Usefulness Assessments

PUR user interface prototype. Table 2 shows the results of the usability test of the PUR user interface. Participants were able to successfully complete all tasks, but at least one error was recorded in every task.

On **task 1** participants were invited to check the details of the rehabilitation session. Task completion time was on average 33.8 s (max 54 s, min 18 s). Three errors were recorded, the first related to the absence of feedback upon entering the login code, leading participants to click multiple times on the screen. Another error was due to the poor visibility of the sessions name. Usefulness was rated around 6. When prompted to provide suggestions, participants proposed that instead of a dark silhouette, the exercises should be exemplified with more appealing and dynamic videos. The participants also stated that the instructions were lacking for more complex exercises.

Task 2 asked participants to change the settings of the rehabilitation session scenario. The task was completed on an average of 26.8 s (max 58 s, min 10 s) and rated high (6) in usefulness. Errors arose from the lack of feedback; i.e. when participants made a selection, no visual confirmation was provided. Options in hidden menus also required a lengthy exploration. Participants suggested the provision of confirmation messages after a button is pressed and asked for more customisation options and the possibility of playing music.

Task 3 prompted participants to perform a rehabilitation session. Average task duration was 40.4 s (max 54 s, min 30 s). Again, this task received very high (6.8). As in task 1, participants had difficulties to find the sessions and it took them a

Table 3 Results of the tests on the RP user interface prototype

Participant	Task 1			Task 2			Task 3			Task 4			Task 5			Task 6		
	D	E	U	D	E	U	D	E	U	D	E	U	D	E	U	D	E	U
RPTP1	90s	1	5	50s	1	6	105s	1	6	80s	1	7	15s	0	6	7s	0	6
RPTP2	10s	0	6	135s	1	7	20s	0	7	50s	1	6	10s	0	7	50s	1	7
RPTP3	55s	2	7	35s	1	6	17s	0	7	80s	0	7	30s	0	7	160s	1	6
RPTP4	3s	0	6	22s	0	5	40s	1	7	90s	1	7	60s	1	7	10s	0	7
RPTP5	7s	0	7	150s	1	7	15s	1	7	45s	0	7	60s	1	7	20s	0	7
Mean	33s	0.6	6.2	78.4s	0.8	6.2	39.4s	0.6	6.8	69s	0.6	6.8	35s	0.4	6.8	49.4s	0.4	6.6

Legend RPTP—rehabilitation professional test participant, D—task duration, E—errors, U—usefulness

bit to get into the exercises. Instructions were unclear at times, which lead to misunderstandings and the incorrect performance of some of the exercises. Since the implementation of a responsive system was not yet developed, some functionalities were misplaced, e.g. the forward button and key images appeared cropped. Participants proposed changing the layout to allow information to be read promptly, without having to search for it on the screen.

Task 4 asked participants to assess the overall session as well as the pain and fatigue experienced post-exercise. Task completion took an average of 13.4 s (max 25 s, min 9 s). Four errors were recorded due to the poor feedback response when selections were made, leading participants to wander through the screen trying to find confirmation on the choice made. Also, some buttons appeared on the user interface, for which the functionality had not yet been implemented leading to confusion. Despite errors, this task was ranked as very useful (6). Similarly to task 2, participants suggested that a confirmation message was provided when a selection was made.

Task 5 prompted participants to leave a message to the RP. The task was rated as very useful (6.4) and took an average of 19.2 s (max 31 s, min 7 s). Due to lack of feedback, participants tried to click multiple times on "record" which slowed down the task. Participants suggested that a "send" button is displayed once they finished recording the message.

Task 6 asked participants to exit the session, as if they were too tired. The task took an average 20.6 s (max 38 s, min 13 s) to complete and was rated very useful (6.4). Participants only had to logout the session; thus, we expected the task to be quickly completed, but this was not the case because of the "*Sair*"[2] button was hidden in the menu and not immediately visible on the screen. All participants suggested that when "*Sair*" button was pressed, the application should return to the main menu, instead of logging out of the application.

RP user interface prototype. Table 3 summarises the results of the RP user interface usability tests. All tasks were successfully completed almost effortlessly.

Task 1 asked participants to check the level of fatigue of their patients. The task was completed on an average time of 33 s (max 90 s, min 3 s); however, three errors

[2] Sair is the Portuguese word for exit.

were recorded. RPTP1 tried to access the information through the patient's individual page, while RPTP3 tried to get that information by accessing the list of all patients, to then check fatigue in each exercise instead of the patient's fatigue. In discussing what might have motivated the errors and proposing suggestions, the participants suggested that the pain and fatigue scales were renamed to EVA and Borg scale, to follow their usual practice. Usefulness was rated useful (5) to extremely useful (7).

Participants took an average time of 78.4 s (max 150 s, min 22 s) to complete **task 2** which aimed to monitor the performance of a PUR. Four errors were recorded, where two participants experienced difficulties identifying the sessions assigned to each PUR, one was unable to locate the information associated with a specific PUR, and another could not find the sessions nor the exercises due to not scrolling to the end of the page. Participants found this task very useful. Participants proposed that session plans expire after a period of time, and that they would like to be able to see the last time the PUR's had completed a session/exercise and to access the PUR's medical history.

Task 3 aimed to assess the fatigue level of a specific exercise and was completed on an average time of 39.4 s (max 105 s, min 15 s). Three errors were recorded, where one participant sought the information on the screen presenting an overview of all PURs. After thoroughly exploring that page, the participant eventually noticed the link *Utentes*.[3] One participant thought effort was associated with the exercise intensity and another needed to read each button label before completing the task. This task ranked as extremely useful. One participant suggested that the exercise details feature should be improved, especially the chart displaying the total number of instructions.

Task 4 asked participants to create a general session plan. All participants completed the task within an average time of 69 s (max 90 s, min 45 s). Three participants tried to use the exercise filter field to characterise the session, instead of searching for exercises categories. The task received a score of 7 on usefulness. One participant suggested the addition of videos demonstrating the exercises and that sessions became unavailable after a period of time.

Task 5 prompted participants to change the characteristics of an exercise, which took an average of 35 s (max 60 s, min 10 s). The two RPs who took the longest, first browsed the general session plans page, instead of the PUR's dedicated page. This task was found to be very useful and no suggestions were made for improvement.

Task 6 asked participants to assign a previously created session plan to a patient and took an average time of 49.4 s (max 160 s, min 7 s). Like the previous task, participants first looked in the general session plans page instead of the page of a specific patient, which was recorded as error. This task was rated as very useful. Participants further suggested that it should be possible to access PURs' sessions history as well as their exercise performance collected data.

[3] Utentes is the Portuguese word for a short version of person undergoing rehabilitation.

Table 4 PUR—system usability scale results

Participant	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	SUS score
PURTP1	4	2	4	1	4	4	4	1	4	1	77.5
PURTP2	5	3	3	1	3	5	3	2	3	3	57.5
PURTP3	3	1	5	1	5	4	5	1	4	1	85
PURTP4	4	1	3	1	3	4	4	1	3	1	72.5
PURTP5	4	2	3	1	4	4	3	2	3	1	67.5

Legend S—statement, PURTP—person undergoing rehabilitation test participant

Table 5 RP—computer system usability questionnaire results

Participant	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	Mean
RPTP1	6	7	7	7	5	4	1	3	7	4	6	5	7	7	3	6	5.3
RPTP2	6	7	6	7	7	7	6	7	7	7	6	7	7	7	7	7	6.8
RPTP3	7	7	6	7	7	7	2	7	7	7	7	7	7	7	7	7	6.6
RPTP4	6	6	7	6	7	7	2	6	6	6	6	7	7	7	7	7	6.3
RPTP5	7	6	6	5	7	7	2	5	7	7	7	7	7	7	6	7	6.3
Mean	6.4	6.6	6.4	6.4	6.6	6.4	2.6	5.6	6.8	6.2	6.4	6.6	7.0	7.0	6.0	6.8	6.2

Legend S—statement, RPTP—rehabilitation professional test participant

3.2 Post-session Self-report Usability Questionnaires

PUR—System Usability Scale. Upon finishing the usability testing session, participants testing the PUR prototype were asked to complete the SUS. Table 4 synthesises the results, where after calculating the scores, results in a score of 72, which means the overall system usability is acceptable.

RP—Computer System Usability Questionnaire. To assess the usability of the RP prototype, we asked participants to answer the Computer System Usability Questionnaire. Table 5 summarises the results. CSUQ results can be analysed in four categories: system utility (statements 1–6), information quality (statements 7–12), interface quality (statements 13–15) and general satisfaction (statements 1–16). The analysis of the results reveals that the category with the lowest average score is information quality. This result (5.7) is affected by the poor score obtained in question 7, which concerns error messages. The average scores obtained in the system utility (6.5) and interface quality (6.7) categories are very positive, where two items in the interface quality category get full scores by all users. The global average is also positive.

Net Promoter Score. The last step of the evaluation assessed the likelihood of recommendation of the system. This was obtained using the Net Promoter Score that asks the question *How likely is it that you would recommend this system, to a friend or colleague?* that is answered using an 11-point scale between 0 (not at all likely) and 10 (extremely likely). In this scale, ratings of 9 or 10 are "promoters", ratings of 7 or 8 are "passives", and rating of 6 or lower are "detractors". The score of the PUR user interface was 80 (PUR1 = 10, PUR2 = 10, PUR3 = 8, PUR4 = 10, PUR5

= 9), where one participant is "passive" and four are "promoters". The score for the RP user interface was 60 (RP1 = 7, RP2 = 9, RP3 = 10, RP4 = 9, RP5 = 10), where two participants are "passives" and three are "promoters". With regard to PUR user interface, all participants said, despite the difficulties experienced during the test, they believed that once glitches were resolved the system was not only welcome but also very much needed.

4 Discussion and Future Work

Despite the occurrence of errors, these were not critical to the point of preventing participants from completing the tasks. Errors did however influence time on task, namely task 6 of the PUR user interface prototype and task 2 of the RP. Although the number of errors recorded for each task of both user interfaces was not particularly high, it is important to analyse each of these errors so that they are adequately resolved and the user interfaces are iterated accordingly. In the PURs user interface, errors were mostly related to poor layout and buttons being mispositioned making them hard to be seen. This made it hard to navigate through the user interface. The comments and suggestions of the participants provided important insight on how the user interfaces could be improved. For example, PUR participants suggested the addition of a button to send the message to the RP once the recording had been completed and the possibility to play music alongside with the sessions. Several errors derived from poor feedback, which confused users and led them to think they were doing something wrong. This will need to be carefully observed in the next iteration to ensure that the system always provides clear feedback for all the actions in the application. Responsiveness also needs to be ensured. The usability tests of the RP user interface were also important to identify areas of improvement, that while minor, were still important to note. Examples include the need of having scales correctly named on the plots and the need to display the medical history and diagnosis of the patient. Participants also suggested the inclusion of a limited time frame for sessions execution and also exercise demo videos. Both user interfaces need also to be reviewed to include pop up confirmation and error messages throughout the application to facilitate the use and prevent errors. In addition to the above corrections, an interactive component that introduces playful elements of gamification is going to be developed to increase adherence, where the RP motivates the PUR to carry out a rehabilitation programme in a personalised way by providing remote monitoring and guidance.

5 Conclusions and Final Remarks

This project aims to create a telerehabilitation platform that allows for real-time monitoring of the movements of PUR and assures that the rehabilitation protocol is carried out correctly. The development of this platform will contribute to improve equity and access to rehabilitation care, with quality and reliability. This paper presented the evaluation of the first user interface prototypes. The evaluation involved ten people, five persons testing the PUR user interface and five physiotherapists testing the RP user interface. Results highlight a number of corrections and improvements that need to be carried out before the rehabilitation platform can be evaluated through pilots in a real context.

Acknowledgements The authors thank the participants of the evaluations. This work was supported by FCT—Foundation for Science and Technology, project CISUC—UID/CEC/00326/2020, and by project CENTRO-01-0247-FEDER-047148 INPACT—"Intelligent Platform for Autonomous Collaborative Telerehabilitation" financed by the Portugal 2020 programme and European Union's structural funds.

References

1. Musculoskeletal Health—World Health Organization. www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions
2. Crawford JO, Berkovic D, Erwin J, Copsey SM, Davis A, Giagloglou E, Yazdani A, Hartvigsen J, Graveling R, Woolf A (2020) Musculoskeletal health in the workplace. Best Pract Res Clin Rheumatol 34(5):101558. https://doi.org/10.1016/j.berh.2020.101558
3. Vegesna A, Tran M, Angelaccio M, Arcona S (2017) Remote patient monitoring via non-invasive digital technologies: a systematic review. Telemed E-Health 23(1):3–17. https://doi.org/10.1089/tmj.2016.0051
4. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS (2017) Telerehabilitation: review of the state-of-the-art and areas of application. JMIR Rehabil Assist Technol 4(2):e7511. https://doi.org/10.2196/rehab.7511
5. Tornero-Quiñones I, Sáez-Padilla J, Espina Díaz A, Abad Robles MT, Sierra Robles Á (2020) Functional ability, frailty and risk of falls in the elderly: relations with Autonomy in daily living. Int J Environ Res Public Health 17(3):1006. https://doi.org/10.3390/ijerph17031006
6. 14:00-17:00 ISO 9241-210:2019. In: ISO. https://www.iso.org/standard/77520.html Accessed 23 Sept 2022
7. Brooke J (1995) SUS: a quick and dirty usability scale. Usability Eval Ind 189
8. Lewis JR (1995) IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J Hum-Comput Interact 7(1):57–78. https://doi.org/10.1080/10447319509526110
9. Albert B, Tullis T (2013) Measuring the user experience: collecting, analyzing, and presenting usability metrics, 2nd edn. Morgan Kaufmann, Amsterdam; Boston
10. Lewis JR (2018) Measuring perceived usability: the CSUQ, SUS, and UMUX. Int J Hum-Comput Interact 34(12):1148–1156. https://doi.org/10.1080/10447318.2017.1418805

Stress Detection and Monitoring Using Wearable IoT and Big Data Analytics

Arnav Gupta, Sujata Joshi, and Menachem Domb

Abstract Stress is a natural response to various stressors that can result in physiological, social, and behavioral changes. Stress can severely affect our bodies if it lasts for a long time. There are numerous things and factors that have an impact on millions of people's lives. As a result, informing the individual about this hazardous lifestyle is critical as warning them before an acute problem develops. The subject's body temperature, heart rate, and galvanic skin response are required to determine stress levels. Wearable IoT-based body sensors give the needed data about an individual cognitive, mental, and emotional health. The paper focuses on utilizing an IoT-based sensor model to identify a person's stress level and deliver feedback to help the individual cope with stress and pressure. An intelligent wristband and chest strap module, placed on the hand and chest, are part of the proposed model. IoT integrated with analytics evaluates activities like electrodermal and heartbeat, further delivering the information to a server that acts as an online IoT cloud-based platform. AWS IoT Analytics integrated with Tableau are used to analyze the data resulting in relevant visualized reports used by the individual to deal with the situation, such as visiting a medical professional or practicing meditation or yoga techniques.

Keywords IoT-based sensor · Galvanic skin response · Electrodermal activity · Heart rate · AWS IoT Analytics

A. Gupta · S. Joshi
Symbiosis International (Deemed) University, Pune, India

M. Domb (✉)
Ashkelon Academy College, 12 Ben Zvi, Ashkelon, Israel
e-mail: dombmnc@edu.aac.ac.il

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_61

1 Introduction

According to medical research, physical, psychological, and behavioral disorders trigger stress. Early detection and treatment of stress can help reduce the symptoms, which further helps reduce fatalities and economic disruption. Thanks to technological advancements, wearable gadgets with physiological sensors have made it easier to track stress. These gadgets detect and decode bio-signals generated by the human body in every day and stressful situations [1]. Connecting these devices to the Internet may collect, share, and store the data in the cloud for analysis by a specialized data decision-making system and be visualized. As a result, remote monitoring and analysis of stress patterns are essential aspects of the stress-reduction strategy [2].

Acute stress and chronic stress are two basic types of stress. Acute stress occurs when the body reacts to a stressful event for a short time before returning to its normal state. The stress that lasts for a long time and has the potential to harm our bodies is called chronic stress. Blood sugar, cholesterol, severe headaches, heart disease, mental health concerns, liver problems, cancer, and other disorders are all linked to stress. Chronic stress can trigger cancerous cells and cause tumor cells to develop more rapidly in people with cancer. At the same time, it also raises the chance of hypertension in cardiac patients, which is unfavorable [3]. Understanding the anxiety rate in patients like cancer patients and cardiac patients might help them recover more quickly. As a result, it is critical to identify a person's stress level well before it begins to have negative consequences on our bodies [4].

Wireless technology has become increasingly important in various industries to deliver improved health care and devices for continuous monitoring. The fundamental motivation for this study is to create a self-stress detection and monitoring system to reduce the harmful effects of stressful conditions on the individual's mental and physical health. Health types are associated with the sympathetic nervous system, which is stimulated during stressful situations, and physiological characteristics, such as electrodermal activity (EDA), which includes temperature and humidity, respiration, heart rate (HR), and blood pressure (BP), are taken into account [5]. The data is recorded and saved on a general storage server in several current systems. This research proposes the use of the AWS IoT Analytics platform. For data reception and storage, AWS IoT Analytics requires an authenticated account. Amazon Web Services (AWS) IoT analytics platform inputs data and runs analysis in real time [6]. The user must have a registered account and a channel to receive data from the microcontroller. Data visualization apps such as Tableau enable the generation of reports on the data processed. The transmitted data is transferred to the channel and then analyzed, allowing stress detection and monitoring.

The paper structure is as follows: Sect. 2 briefly discusses the research gap. Section 3 gives the objective of the paper. Section 4 describes the literature review, giving an overview of the previously performed related work. Section 5 discusses the research methodology used in this paper. Section 6 discusses the research questions the paper tries to address. Section 7 gives the system overview describing various concepts like physiological parameters, system architecture, and software–hardware details. Section 8 describes the conclusion of the research. Section 9 gives the research implications, and Sect. 10 briefs the possible future scope of work.

2 Research Gap

The IoT sensors used in stress detection and monitoring generate massive data volume. Continuous data collection creates enormous amounts of information that must be collected and analyzed. This data must be accurate, reliable, complete, and appropriately represented to support decision-making. This paper addresses the stress monitoring and analysis gap using specified wearable sensors and big data analytics techniques.

The IoT wearable sensor devices for monitoring and detecting stress satisfy the conditions to provide high-quality, reliable data with timestamps for further insights. Timestamps allow data synchronization. Big data analytics provide actionable insights. Choosing the correct significant set of information from the data gathered by the sensors plays a crucial role in monitoring stress.

3 Literature Review

An abundance of stress occurs due to various emotional, social, and even physical manifestations; as a result, the pressure's adverse effects vary significantly across people. People with excessive stress experience sleeping disorders, migraines, muscle strains, weakness, and even metabolic problems. Anxiety, instability, changes in food patterns, loss of excitement or energy, and mentality are some emotional and behavioral side effects of stress.

3.1 IoT-Based Sensors for Stress Monitoring and Detection

Sensors are becoming increasingly significant in healthcare-related industries. A wearable sensor is an electronic device that uses one or more sensors, such as BP, GSR, and HR sensors [7]. Stress is often acknowledged as one of the key contributors to various health conditions that, if left untreated, can be fatal. Ogorevc et al. examined how cognitive stress affected particular psychophysiological variables by analyzing emotional and physical stress based on individual tasks. When people are exposed to stressful activities, their blood pressure, heart rate, and galvanic skin response levels increase [8].

3.2 Usage of Data Analytics Techniques in IoT-Based Sensors

Big data analytics is a big step forward in storing and processing vast amounts of data efficiently, allowing for creativity and innovation about how to use the results in a meaningful and helpful way. Its unique attributes will open new perspectives on how data analytics in healthcare promotes public health at a low cost [9]. AWS IoT analytics software enables us to connect to various devices, then access, and process their data. We can then use the knowledge gathered from these stages to build an automated system to regulate the sensors and generate meaningful output. Dineshkumar and Senthilkumar proposed the use of big data analytics in healthcare monitoring systems using the Hadoop framework, which allows the model to perform real-time monitoring to alert the patient [10].

3.3 Various Studies in This Area

This section describes previous research on stress evaluation, categorization, and application in evaluating multiple medical problems. Respiratory rate (BPM), body temperature, heart rate (HR), blood pressure (BP), galvanic skin response (GSR), and other physiological measures are used for assessing people's response to stress [11]. Basel Khikia et al. created a hand band for individuals with mental illness that included a galvanic skin response sensor and high beam sensors to categorize "Stressed" and "Not stressed" incidents. They conducted analyses on individuals as personnel observed their cognitive and emotional patterns. After analyzing data from the sensors, they discovered that various stress level situations handle multiple situations [12]. Seoane et al. set out to create a wearable gadget that would allow warriors to assess their physiological, emotional, and cognitive stress levels during the fighting. They divided their project into two phases, accomplishing the first step in this publication. The first phase determined the best biomedical parameter for analyzing stress. In contrast, the second phase was dedicated to developing sensor-based wearable metrics for monitoring and analyzing the biomedical parameter to achieve various stress levels experienced by combatants [13]. Because essential characteristics can be obtained from ECG, like heart rate (HR) and respiration rate, which is essentially self-working by the nervous system, they successfully discovered that ECG is the ideal bio-signal to measure a person's state of mind. They plan to develop a wearable device that measures and analyzes ECG to detect stress in the long term [14].

The physiological sensing-based stress analysis assessment was investigated as part of an examination of the study on the usefulness of GSR in stressed conditions. Several different questions with varying levels of difficulty are provided to students with GSR sensors worn by them for evaluation. The data shows a substantial correlation between GSR measurement and the stress brought on the brutal questioning [15]. The accelerometer sensor, in combination with the GSR sensor, can be utilized to improve stress analysis. Combining activity awareness with stress detection utilizing a GSR sensor enhances the reliability of the stress evaluation [16]. The stress states of the drivers are evaluated using GSR sensor readings and accelerometer data, and they are classified as stressed or non-stressed. Different projects also concentrated on creating, building, and developing a GSR sensor to increase effectiveness and efficiency and lower mistakes [17].

4 Research Methodology

The qualitative research methodology is used in this research. The keywords are "mental stress detection," "mental stress," "IoT-based sensors for stress," "analytics for stress detection," "monitoring of stress," and "stress detection using sensors." This research aims to provide a detailed overview of how stress can be monitored and detected effectively using sensors and big data analytics. This paper proposes a methodology using an IoT sensor measuring EDA, HR, and BP to measure and monitor stress accurately. Smartwatch and chest strap contains IoT sensors connected to smart devices, allowing them to be used as a resource for physiological parameters and quality processing of physical and mental health in stress detection and monitoring. This paper addresses the requirement:

(a) Using an IoT-based wrist and chest sensor as a source for physiological parameters such as heart rate (HR), blood pressure (BP), and body temperature.
(b) Identifying the individual's activities by analyzing the data using the AWS IoT analytics platform from the IoT-based wrist and chest sensor.
(c) Providing real-time feedback using contextual information by data analysis.

5 Research Questions

For this research, we used a case study technique, in which numerous use cases of intelligent health monitoring system and their benefits to diverse sectors were evaluated and debated. This study's information came from various online databases, whitepapers, publications, and reports. The following are the research questions addressed in this study:

1. Does the IoT-based sensors in use with big data analytics and automation play an essential role in improving operating efficiency and service delivery in the healthcare industry?
2. Can actionable insights be derived after using the proposed system to help the hospitals or the users?
3. Can the proposed model be helpful to the general public and hospital patients to improve their emotional and cognitive state of mind?

6 Data Model and Findings

This section discusses system architecture, software and hardware details, and the parameters it depends on.

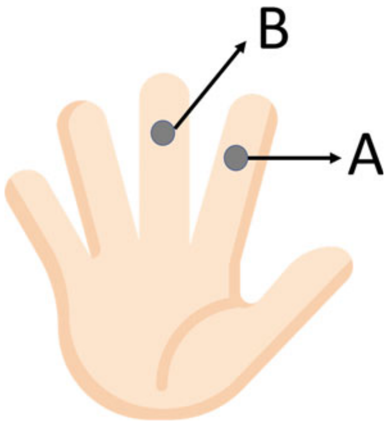
6.1 Physiological Parameters

The activity of the sympathetic nervous system can be monitored using a variety of functional measures such as blood pressure (BP), heartbeat rate (HR), respiratory rate, the temperature of the body, electroencephalogram, and the galvanic skin response (GSR). Electrodermal activity (EDA) and heart rate (HR) are incorporated in this suggested model since these characteristics can be evaluated with an IoT-based sensor model and are connected to the human central nervous system. Physiological parameters are used as a reference value to measure any matter related to human health monitoring. With these parameters, one can determine how an individual acts in a particular state. In this paper, these parameters can actively determine the body response that an individual may have during stressful activities.

The galvanic skin response (GSR): The sweat glands are significant to our bodies. There are more than millions of sweat glands in the human body. The hands, forehead, and feet contain the majority of them. Our sympathetic nervous system is directly linked to our sweat glands. The sympathetic nervous system is a part of the human nervous system that engages when we are agitated, such as when we are stressed [18]. Sweat travels through the pores of our skin when the sympathetic nervous system triggers the sweat glands. As a result of the produced fluid, the epidermis becomes wet. This fluid contains positively and negatively charged ions, altering our skin's electrical resistance. Galvanic skin response represents a shift in change in skin resistance caused by emotional situations. (GSR) [19]. Grove GSR is the galvanic skin response sensor utilized in this study. It comprises two electrodes that monitor the subject's electrical resistance and provide the desired voltage. The device is frequently put on emotionally vulnerable body parts like our hands, fingers, and soles of our feet. Figure 1 depicts the GSR measurement site.

The position of the sensor should be on the index and middle fingers of the individual's non-dominant hand, designated as *A* and *B* in Fig. 1. A GSR signal is made of two primary components. These are the "tonic" and "phasic" components. The tonic component is a response that fluctuates minimally and slowly; therefore, the name "slow component." In contrast, the phasic element is a response that displays considerable variations and is evident as "GSR peaks," so the name "fast component." The latency (delay), recovery time, rise time, and peak amplitude are the characteristics of the GSR signal [20]. Stimulus onset refers to the time when stimulation is administered to a person. Latency is the time interval between the stimulation's start and the GSR's apex. Peak amplitude is the disparity between the strength of the GSR signal at stimulation commencement and the intensity of the GSR peak [21]. Once

Fig. 1 Measurement of GSR



stimulation has begun, the rise time refers to the time it takes for a GSR signal to attain its maximum value. The recovery time it takes for the GSR response to return to its stimulation balance after reaching a peak is the recovery time [22].

Heart rate (HR): The heart is a blood-pumping machine essential to human survival. It operates as a circulatory pump for blood in our body's functions. The body's autonomous neurological system controls the heart's function. The body requires extra oxygen for energy build-up in a "fight or flight" situation [23]. Blood is a type of connective tissue that transports oxygen throughout our bodies. As a result, whenever our bodies want more oxygen, our bodies' independent neurological systems urge the heart to pump additional blood into the artery that distributes oxygen throughout the human body, raising the heartbeat [24]. Heart rate is the rate at which the heart circulates blood into an artery at a specific time. As a result, we can deduce that our autonomous nervous system is inextricably linked to our heart rate [25].

6.2 System Architecture

The prototype's system architecture comprises two sets of sensors: a smart band on the wrist and a strap worn around the chest. The setup model measures the electrodermal activity, blood pressure, and heart rate of the subject, i.e., the individual whose stress level is being assessed. Figure 2 depicts the proposed work's system design. An open IoT platform, AWS IoT Analytics, integrated with business intelligence tools such as Tableau, is used in the system's overall architecture. Sensing elements, microcontrollers, and communication design features are the essential components. Sensing devices detect various physiological characteristics and are then supplied to the microcontroller. Because the information from the sensors is unprocessed, the microcontroller will employ signal processing techniques like filtering and sampling

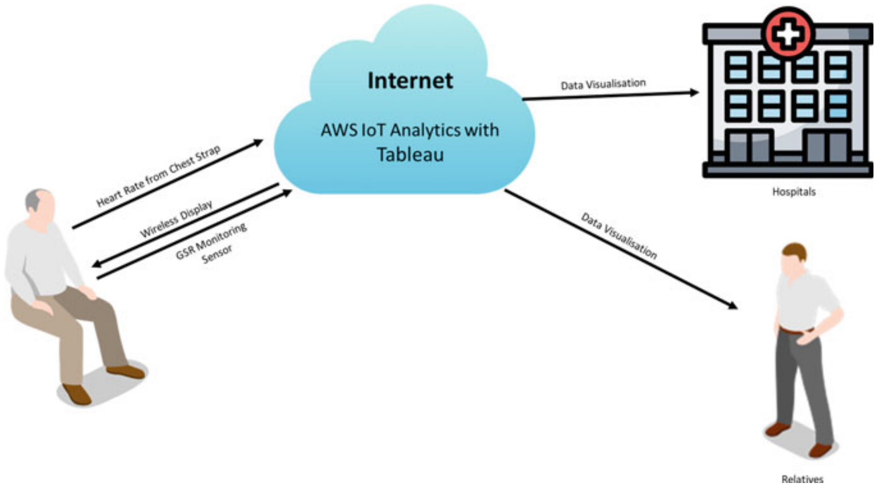


Fig. 2 System design

before forwarding it to the site. The connectivity components can transfer data from the microcontroller to the open Internet cloud-based system.

Users can do online computations on the information they obtain using the open Internet system. An account user must examine the components required for the information delivery from the microcontroller to access data. A microcontroller and a wireless functioning module enable a TCP connection for data transfer, allowing the sensors to communicate with these channels. The unprocessed raw information from the sensors is transmitted to this cloud infrastructure after the link is formed, where it is later processed and evaluated. Big data analytics is used to quantify the occurrence of stressful events the user has experienced. The frequency of stressful circumstances is displayed on the dashboard and charts to visualize data.

6.3 Software and Hardware Details

The system consists of software and hardware components that perform two main functions: collecting data communicated by sensors and analyzing data collected by AWS IoT analytics with Tableau.

Arduino IDE—Arduino is a free, open-source microcontroller that can be coded, cleaned, and reconfigured anytime. The platform was designed to make it easy and affordable for people to build gadgets interacting with their environment using sensor devices. It is an open-source framework for creating and managing electronic gadgets built around low-cost microcontroller boards. It accepts inputs from various electronic devices and regulates their outputs. The data is collected using the "analogRead" function on the Arduino board. The function analogRead can read the value from the specified pin. It takes 100 microseconds to read an analog input. Arduino ESP32 can be used to make connections between the AWS platform. The Wi-Fi SSID and host configuration should be made, and the platform can be ready to use.

AWS IoT Analytics—AWS IoT Analytics is a robust tool that uses simple analytical models for massive amounts of IoT data and executes them successfully. Amazon Athena is a Web-based cloud storage tool that allows professionals to run dynamic searches in Amazon Simple Storage Service (S3). Large amounts of data are employed with Athena. On Amazon Web Services, Amazon S3 is used for online storage and preservation of information assets. With use cases like data processing, recordkeeping, Web site management, information backup and restore, and project hosting for deployment, Amazon S3 was intended to make online processing easier for programmers. Amazon Athena allows customers to utilize structured query language to analyze the data in Amazon S3 (SQL). The Athena connector can be connected to Tableau to create dashboards from the data received by sensors.

Tableau is a rapidly evolving business intelligence (BI) visualization application. The Tableau Big Data platform allows you to capture, analyze, and handle more information than ever. Tableau's analytics platform enhances IT and includes security features, regulation, installation flexibility, and administration. Tableau Analytics enables businesses to get much more out of their data and workforce. Tableau offers a variety of analytics tools, like Desktop, Prep Builder, and Tableau Server. This visual analytics platform can transform unprocessed raw data into actionable insights and solves problems by visualizing important information. Data that is significantly important to the user can be imagined. Insights are derived from the dashboard, and the user can take action.

Sensors—The hardware design comprises two units, an intelligent wristband unit and a chest strap unit used to measure physiological data at two locations. The intelligent wristband module, placed around the wrist, comprises a galvanic skin response (GSR) sensor, an Arduino LilyPad microcontroller with a voltage regulator, a power supply, and a Wi-Fi module. Figure 3 depicts the model's architecture containing the IoT-based chest strap and wrist sensors attached to the Wi-Fi module. The Arduino LilyPad microcontroller gets data from a GSR sensor that monitors electrodermal activity (EDA) and is connected to the Arduino board. The ECG module in the chest strap module detects the heart's electrical activity with a three-lead ECG electrode. The devices collect the information and transfer it to the microcontroller. A GSR sensor, Arduino LilyPad power supply, Arduino LilyPad microcontroller, voltage regulator, and a Wi-Fi module make up the smart band module, which is worn on the wrist. The Arduino LilyPad processor receives data from a GSR sensor linked to the Arduino board, which monitors electrodermal activity. The data is then delivered to the AWS IoT Analytics software over a TCP connection established with the help of the Wi-Fi module. The Wi-Fi module uses the AT instructions to communicate with the IoT platform. The ECG module with a three-lead ECG electrode in the chest strap module detects the heart's electrical activity. The data is collected by the module and transmitted to the microcontroller performing additional signal processing on the data, such as lowpass filter screening and sequencing. The heart

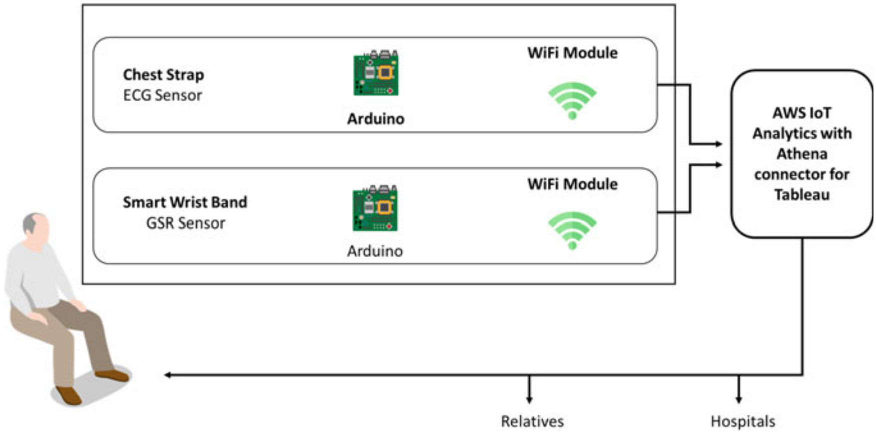


Fig. 3 Overview architecture of the proposed model

rate is computed using preprocessed data. The heart rate data is then transferred to the AWS IoT platform using the Wi-Fi module to establish a connection.

7 Conclusion

In the healthcare industry, wearable sensor devices combined with IoT technologies are significant. The advantages of employing such devices have benefited both patients and clinicians. Human stress detection is critical because excessive stress can affect a person. This paper overviews the monitoring and detection of stress with IoT wearable sensors and big data analytics. A qualitative research methodology is used, and a model is proposed using the AWS IoT platform having a Wi-Fi module and wearable sensors like wristbands and chest straps. The data from the sensors is passed on to the cloud-based AWS IoT platform through the Wi-Fi module. The data can then be visualized by Tableau for a better understanding by continuously monitoring and delivering ongoing feedback to the user. The model will increase the efficiency and effectiveness of existing models by providing real-time feedback and, in turn, help deliver better health aid to a person.

8 Research Implications

Stress is a heightened psychophysiological state of the human body that occurs in response to a stressful event or situation. Prolonged stress can cause a lot of emotional and mental problems for humanity. IoT-based sensors and big data analytics techniques can detect and monitor stress. The data from the sensors can be captured and then can be analyzed. This model can be beneficial not only for clinics and hospitals but also for the general public. The model can be used daily to explore a particular individual's trend. Further actions such as meditation and yoga can mitigate stress levels.

9 Research Limitations

There are limitations to the research that can be addressed in the future scope of work. Firstly, the sensors should be placed correctly to minimize ambiguities and adhere to criteria for accurately collecting physiological parameters. Device noise, random noise, loose instrument skin connections, and physical movement affect signals.

Secondly, stress measurement instruments should be non-invasive to acquire valid data. When invasive devices are worn, they might add to the individual's stress. Smart wearable systems are frequently used to capture large volumes of data discreetly, sometimes without the user's knowledge.

10 Future Scope

The future scope of work on this research can be focused on the following. Firstly, combining machine learning and deep learning with big data analytics makes the model more effective in monitoring stress. Secondly, use new inbuilt smart IoT sensors that can be continuous and accurate data assessment by collection and analysis without the cloud system to show real-time stress monitoring. This model can also detect stress in students, instructors, and corporate and office workers.

References

1. Gedam S, Paul S (2021) A review on mental stress detection using wearable sensors and machine learning techniques. IEEE Access 9:84045–84066. https://doi.org/10.1109/ACCESS.2021.3085502
2. Can YS, Arnrich B, Ersoy C (2019) Stress detection in daily life scenarios using smartphones and wearable sensors: a survey. J Biomed Inform 92:103139. ISSN 1532-0464. https://doi.org/10.1016/j.jbi.2019.103139
3. Schneiderman N, Ironson G, Siegel SD (2005) Stress and health: psychological, behavioral, and biological determinants. Annu Rev Clin Psychol 1:607–628. https://doi.org/10.1146/annurev.clinpsy.1.102803.144141. PMID: 17716101; PMCID: PMC2568977
4. Subramanya K, Bhat VV, Kamath S (2013) A wearable device for monitoring galvanic skin response to predict changes in blood pressure indexes accurately and cardiovascular dynamics. In: India Conference (INDICON). IEEE, pp 165–168
5. Widanti N, B. Sumanto, P. Rosa and M. Fathur Miftahudin, "Stress level detection using heart rate, blood pressure, and GSR and stress therapy by utilizing infrared," *2015 International*

Conference on Industrial Instrumentation and Control (ICIC), 2015, pp. 275–279, DOI: https://doi.org/10.1109/IIC.2015.7150752.
6. Marino CA, Chinelato F, Marufuzzaman M (2022) AWS IoT analytics platform for microgrid operation management. Comput Ind Eng 170:108331. ISSN 0360-8352. https://doi.org/10.1016/j.cie.2022.108331
7. Yashaswini DK, Bhat SS, Sahana YS, Shama Adiga MS, Dhanya SG (2019) Stress detection using deep learning and IoT. Int J Res Eng Sci Manage
8. Ogorevc J, Podlesek A, Geršak G, Drnovšek J (2011) The effect of mental stress on psychophysiological parameters. In: Proceedings of IEEE international symposium on medical measurements and applications. Bari, Italy, May 2011, pp 294–299
9. Raghupathi W, Raghupathi V (2014) Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2(1):1–10
10. Dineshkumar P, Senthilkumar R, Sujatha K, Ponmagal RS, Rajavarman VS (2016) Big data analytics of IoT based Health care monitoring system. In: 2016 IEEE Uttar Pradesh Section international conference on electrical, computer, and electronics engineering (UPCON), pp 55–60. https://doi.org/10.1109/UPCON.2016.7894624
11. Fernandes A, Helawar R, Lokesh R, Tari T, Shahapurkar AV (2014) Determination of stress using blood pressure and galvanic skin response. In: International conference on communication and network technologies (ICCNT). IEEE, pp 165–168
12. Kikhia B et al (2016) Utilizing a wristband sensor to measure the stress level for people with Dementia. Sensors
13. Seoane F, Mohino-Herranz I, Ferreira J, Alvarez L, Buendia R, Ayllon D, Llerena C, Gil-Pita R (2014) Wearable biomedical measurement systems for assessment of mental stress of combatants in real time. Sensors 7120–7141
14. Lee B-G, Chung W-Y (2017) Wearable glove-type driver stress detection using a motion sensor. IEEE Trans Intell Transp Syst 18(7):1835–1844. https://doi.org/10.1109/TITS.2016.2617881
15. Sandulescu V, Andrews S, Ellis D, Bellotto N, Mozos O (2015) Stress detection using wearable physiological sensors, pp 526–532. https://doi.org/10.1007/978-3-319-18914-7_55
16. Tang TB, Yeo LW, Hui Lau DJ (2014) Activity awareness can improve continuous stress detection in galvanic skin response. Sensors 1980–1983
17. Das P, Das A, Tibarewala DN, Khasnobish A (2016) Design and development of portable galvanic skin response acquisition and analysis system. In: International conference on intelligent control power and instrumentation (ICICPI). IEEE, pp 127–139
18. Dibona GF (2013) Sympathetic nervous system and hypertension. Hypertension 61(3):556–560
19. Kim D-S, Hwang T-H, Song J, Park S, Park J, Yoo E-S, Lee N-K, Park J-S (2016) Design and fabrication of smart band module for measurement of temperature and GSR (galvanic skin response) from human body. Proc Eng 168:1577–1580. https://doi.org/10.1016/j.proeng.2016.11.464
20. Safa M, Pandian A (2021) Applying machine learning algorithm to sensor coupled IoT devices in prediction of cardiac stress—an integrated approach. Mater Today Proc. ISSN 2214-7853
21. Sinha A, Das P, Gavas R, Chatterjee D, Saha SK (2016) Physiological sensing-based stress analysis during the assessment. In: Frontiers in education conference (FIE). IEEE, pp 1–8
22. Lockhart RA (1972) Interrelations between amplitude, latency, rise time, and the Edelberg recovery measure of the galvanic skin response. Psychophysiology 9(4):437–442
23. Ahuja ND, Agarwal AK, Mahajan NM, Mehta NH, Kapadia HN (2003) GSR and HRV: its application in clinical diagnosis. In: 16th Symposium on computer-based medical systems. IEEE, pp 279–283
24. Kim H-G et al (2018) Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry Invest 15(3):235
25. Pandey PS (2017) Machine learning and IoT for prediction and detection of stress. In: 2017 17th International conference on computational science and its applications (ICCSA), pp 1–5. https://doi.org/10.1109/ICCSA.2017.8000018

Comparing Mixed Reality Hand Gestures to Artificial Instruction Means for Small Target Objects

Lukas Walker, Joy Gisler, Kordian Caplazi, Valentin Holzwarth, Christian Hirt, and Andreas Kunz

Abstract Hand gestures are a valuable means for the instruction of complex handling processes. They are used and perceived in an intuitive way and outperform artificial representations such as arrows or symbols. On the other hand, referring finger gestures require a certain object size to avoid ambiguities, and often they are replaced by artificial means. However, this comes to the cost of reduced intuition due to the change of a hand gesture to an artificial gesture, which consequently makes it more difficult to learn long instruction sequences and keep them in mind. This paper thus introduces study results showing that hand pointing gestures perform well even for small objects, so that unnecessary switches to artificial representations can be avoided in the future.

Keywords Augmented reality · Gestures · Visualization

1 Introduction

When explaining complex manual operation tasks, gestures play an important role together with speech. In this context, hand gestures are considered particularly important [3]. Hand gestures for handling objects are unique since they constrain and predefine the gesture [2]. Aigner et al. [2] classify occurring gestures into pointing, semaphoric, pantomimic, iconic, and manipulative gestures, from which only the pointing gesture is not constrained by geometry.

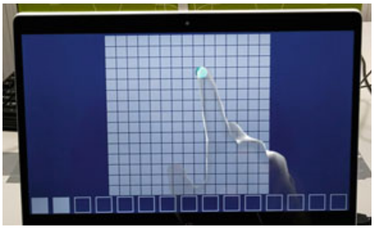
L. Walker · J. Gisler · C. Hirt · A. Kunz (✉)
ETH Zurich, 8092 Zurich, Switzerland
e-mail: kunz@iwf.mavt.ethz.ch

J. Gisler
e-mail: lukas.walker@inf.ethz.ch
URL: http://www.ethz.ch

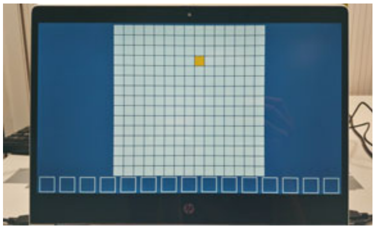
K. Caplazi
Rimon Technologies GmbH, 8092 Zurich, Switzerland

V. Holzwarth
RhySearch, 9471 Buchs, Switzerland

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_62



(a) Fingerpointing overlay displayed by Microsoft HoloLens II



(b) Highlighting the element displayed by laptop screen

Fig. 1 Overview on the two study conditions

If hand gestures are used for an instruction on how real objects should be handled, Mixed Reality (MR) can be used. MR technologies, especially head mounted displays (HMD), receive a lot of attention since a few years. Their success lies in enabling users to interact with digital data seamlessly, which is a great support in both training and in productive settings. The field of MR is subject to many research projects, both in academia and industry, being driven by the question how we can best profit from the new ways to interact with the digital world. Typically, MR superimposes virtual instructions on real objects, as shown in an early work by Thomas and David [25]. This study was continued three decades later by Hoover et al. [9]. The study had different groups performing the task using a tablet with a 2D guide, a tablet with an MR guide, a desktop computer and a HoloLens 1, respectively. The study showed that participants who used the Hololens guide made fewer errors and had faster assembly times by 15%. The study concluded that HMDs like the HoloLens can be a better alternative for state-of-the-art approaches. Also the work from Scurati et al. [23] uses abstract virtual instructions such as arrows or symbols to inform the worker about the next step. This is also confirmed by a systematic review by Palmarini et al. [18], who only listed MR systems that use abstract virtual augmentation. Consequently, it was stated by LaViola et al. [13] that for presenting positions in the real environment, pointing arrows are superior to other virtual overlays. This might be due to the fact that hand tracking was computationally expensive in the past and just recently became available in devices such as the Microsoft HoloLens (I+II) for MR and Oculus Quest (I+II) for Virtual Reality (VR).

In fact, using hand gestures as an instruction means initially required a larger technical effort, as shown in *SEMarbeta* [6], or in *Augmented 3D hands* [10]. Just recently, with the advance of HoloLens, hand gestures could be easily used as an instructive overlay, as shown in [17], who used pointing gestures to emphasize certain objects during the support by a remote expert. Using hand gestures as an overlay for augmented work instructions is a promising approach since humans have an easier access to images, such as gestures rather than to abstract information, such as icons or pictograms. The latter need to be interpreted by a user in order for him or her to understand and perform a task. A hand animation on the other hand shows the task to be performed in real-time and in three dimensions, allowing the user to imitate the movements and view them from different sides. In other words, hand animations show to the user immediately how to complete a task instead of explaining it to him or her, which increases the level of immersion and eventually supports memorization.

Showing hand gestures as an overlay on a manual assembly task will lead to the user mimicking these gestures and thus learning the correct behavior in an intuitive way. However, little is known about the efficiency of such hand gesture overlays in contrast to iconographic instructions particularly when it comes to smaller object sizes. It seems that the accuracy and clearness of hand gestures is limited, since [24] state that hand gestures could also be replaced by virtual pointing rays. Based on literature, there is no clear evidence on the accuracy of hand gesture overlays when it comes to pointing to smaller parts in the real environment. Further, existing MR applications using MR hand gesture overlays focus mainly on a remote support and do not give information on memorization effects which could become relevant in training scenarios. This could help in particular for memorizing longer sequences of numbers, since not the numbers are kept in mind, but the complete fingerpointing gesture which is then related to the underlying matrix for recalling the numbers.

We hypothesize that there is a certain size limit for small neighboring target objects, where a user cannot clearly distinguish anymore, where the fingerpointing gesture is directing them at. Further, little is known on the memorization of sequences that are indicated by pointing gestures in comparison to a regular highlighting using virtual objects. Since Liu et al. [16] describe the positive effect of pointing gestures on spatial memorization, and Aldugom et al. [4] describe gestures' positive effect in learning mathematics, we secondly hypothesize that pointing will positively effect the memorization of work sequences.

This paper focuses on one of the most important gestures—the pointing gesture. We compare this pointing gesture to a common highlighting of relevant positions. The overall goal is to completely use gesture-based instructions, without switching back and forth between gesture overlays and iconographic overlays. This would also reduce development efforts of MR and VR applications that direct a user to certain targets, since hand gestures can be recorded by the device, whereas iconographic overlays have to be manually implemented.

After an overview on pointing possibilities in MR, the paper introduces the user study, which evaluates the limits in accuracy for a pointing gesture, as well as the impact of natural pointing gestures on memorization. This is followed by an evaluation of the achieved data regarding accuracy and memorization. The remainder of the paper will give a brief summary and an outlook on future work.

2 Related Work

Pointing gestures in MR or VR are mainly done using a supporting ray that can be controlled by the hand and the index finger, e.g., for selecting objects as shown by Yusof et al. [27], or for controlling a highlighter as described by Lin et al. [15]. The latter inspired us for our user study, in which also certain fields of a matrix should be

Table 1 Previous works' target sizes

Literature	Target size
Tsang et al. [26]	20
Park et al. [19]	10
Gao and Sun [7]	15.9 × 9
Komine and Nakanishi [12]	7
Leitão and Silva [14]	14
Schedlbauer [22]	15

selected. Highlighters or controlled rays only allow for a remote interaction and thus contain a certain amount of unnaturalness when selecting objects. We therefore use the "direct touch" as being used e.g., by Kervegaut et al. [11] as an inherent HoloLens functionality. For providing feedback when touching virtual objects in MR, handheld devices such as smartphones or tablets are used, as described by Prilla et al. [20], who compared this also to hands-free interaction.

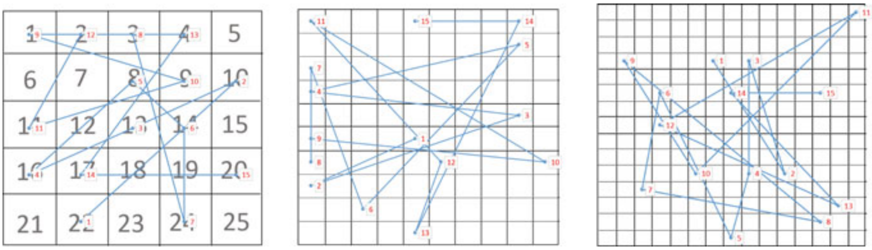
3 Study Design and System Setup

The study described in this paragraph has the following purpose:

- Determine the minimum target size that can be unequivocally detected using a pointing instruction as opposed to a highlighting of objects.
- Effect of a pointing gesture's naturalness on memorizing a sequence of numbers.

To answer these questions, a study was designed where a user was instructed to touch buttons of a matrix on a tablet screen. The target sizes were based on existing manifold works listed in Table 1.

The study consists of three matrices with 5 × 5, 10 × 10, and 15 × 15 elements. Given the resolution of the laptop screen, this results in button sizes of 29 × 29 mm, 14 × 14 mm, and 9 × 9 mm. For all matrices, there was a 0.5 mm spacing between the buttons. This spacing can be seen as irrelevant for the results [22]. For each matrix, the user was instructed which button to press either by a fingerpointing overlay or by highlighting the elements (Fig. 1). For each matrix size, the user had to press 15 buttons. In both cases—the HoloLens instruction and the highlighting—the user's input was detected by the touch sensitive overlay of the laptop computer. The system waits until the user performed the instructed pointing gestures and measures the time for the gesture. As soon as the user touched a button on the screen, the next instruction follows. The instructions where such that the user had to traverse the matrices in seemingly random order (Fig. 2). To avoid any biasing effects due to memorization, every trial of the user study used a different sequence of numbers,



(a) 5×5 matrix (b) 10×10 matrix (c) 15×15 matrix

Fig. 2 Pointing trajectories on the three different matrices

Fig. 3 User study in front of the laptop, showing a 15 × 15 grid for fingerpointing instruction. The lower bar shows the amount of touch inputs already entered

and all of them were randomly chosen so that the generated trajectory cannot be kept in mind.

In a next part of the study, the user was instructed to memorize a sequence of eight buttons that were either highlighted or shown by a fingerpointing overlay in a 4×4 matrix (button size 29×29 mm). The highlighting and fingerpointing were automized, showing a new button every two seconds. After having seen the sequence, the user had to reenter the memorized sequence into the system.

The complete study was designed as a within-subject study, i.e., after signing a consent form, filling out initial questionnaires, and becoming acquainted with the setup, each subject had to do both, the highlighting and the fingerpointing instruction. In order to balance the study, all participants were divided into two groups, which differ in the order of the two initial experiments (fingerpointing and highlighting). The whole study took about 25 min. In order to avoid any technical biasing, users had to wear the HoloLens in both trials of the study, although it was not required for the highlighting task. For the fingerpointing overlay, MS HoloLens II was used, while the highlighting was displayed on a laptop screen (HP ProBook x360 435 G8). The touch sensitive screen of the laptop was used to detect the user's input (Fig. 3).

During the study, three questionnaires were filled out. The first questionnaire collected demographic data as well as a computer confidence level. After each element of the study, the NASA TLX [8] (scale: 1–10) and the SUS [5] (scale = 1–100) questionnaires were filled out by the participants. In addition, the cognitive absorption (CA) [1] was completed (scale: 1–10). Finally, after completing both parts of the user study, a user preference was questioned. 11 participants recruited from the local university staff with an average age of 29 years (SD = 6.05) took part in our user study (2 female, 9 male), from which all had normal or corrected to normal vision.

4 Study Results

4.1 Comparison to Fitts' Law

To make the setup comparable with other works on touch interaction, the three matrices are characterized by measures from the Fitts' law: Index of difficulty (ID), performance index (IP), and motion time (MT). D is the mean length of the paths between the individual touch points in Fig. 2, and w is the size of the buttons in the matrix. The results are summarized in Table 2.

$$\mathrm{ID} = \log_2\left(\frac{2D}{w}\right); \qquad \mathrm{IP} = \left(\frac{\mathrm{ID}}{MT}\right) \tag{1}$$

A comparison of the tables shows that for both—the HoloLens instruction and the highlighting instruction—the index of difficulty (ID) slightly increased with a decreasing button size w, while the traveling distance for a pointing gestures was tried to be kept constant. However, the mean time for performing an action was significantly smaller for the highlighting condition than for the HoloLens, which is also reflected in the performance index IP, which is more than two times higher. The main reason for this reduced performance index when using the Hololens with fingerpointing gestures is that users wait until the fingerpointing gesture is starting to move away from the button, and then start doing the fingerpointing by themselves. Instead of a hand and finger that simply pop up at the correct position (which would be similar to regular highlighting), we consciously chose a moving hand and thus a resulting lower IP, since we believe that a natural movement of a hand is crucial for intuitiveness and more efficient memorization. Consequently, such a "waiting" behavior was not observed for the highlighting condition, and it shows in principle that the pointing gesture overlay is perceived as natural, so that there cannot be "two fingers at the same location".

Table 2 Fitts law measures for holoLens and highlighting

		HoloLens				Highlighting			
Matrix size	w (mm)	D (mm)	ID	IP	MT	D (mm)	ID	IP	MT
5 × 5	29	86.32	2.57	1.07	2.42	84.76	2.55	2.76	0.92
10 × 10	14	93.02	3.73	1.77	2.12	81.21	3.53	3.65	0.97
15 × 15	9	85.04	4.24	159	2.67	57.99	3.69	3.59	1.03

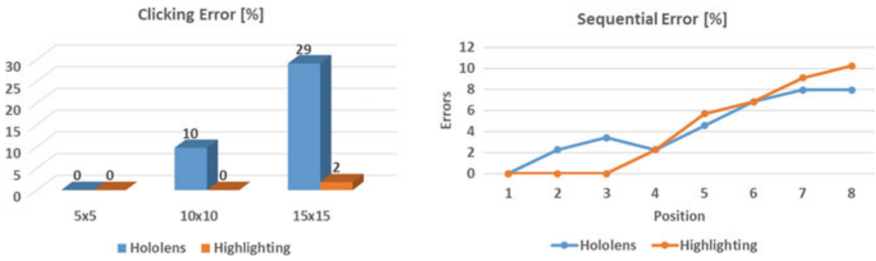


Fig. 4 Cumulative error for HoloLens and highlighting instruction (left); Accumulated errors of the memorization task (right)

4.2 *Evaluation of the Clicking Accuracy*

For all three matrix sizes, the user was instructed to press the button shown to him either by fingerpointing overlay using the HoloLens or by highlighting directly on the laptop screen. In all cases, the error is defined as the sum of wrongly pressed buttons. It was not possible to press two buttons simultaneously or to press a gap between the buttons. The results are shown in Fig. 4 (left). The results show that both—the HoloLens and the highlighting—perform equally well for the 5×5 matrix, while for the larger matrix sizes the amount of wrongly pressed buttons significantly increases for the fingerpointing instruction compared to the highlighting instruction. This is mainly due to the fact that the fingerpointing gesture could not unequivocally be assigned to a button anymore, since it was completely or partially occluded by the finger. For the 10×10 matrix, occlusion was one of the reasons for the occurring errors, since an accidental mistyping due to button size can be excluded as there are no errors for the highlighting instruction. For the 15 × 15 matrix, errors also occur for the highlighting method, which are probably due to accidental mistyping because of the small button size. However, also here the majority of the errors is likely from the wrongly detected buttons due to occlusion by the fingerpointing overlay. Since the forearm, the hand, and the fingers were semi-transparent (Fig. 1a), it was mainly the opaque shape attached to the fingertip that caused the occlusions. However, this shape was necessary to clearly detect the fingertip.

For the memorization study, a 4 × 4 matrix was used to avoid biasing of the results by erroneous readouts due to the small button size. The user had to keep in mind 8 numbers in the right order being shown to him by either fingerpointing or highlighting. Once the instruction sequence was finished, the user had to reenter this sequence on the touchscreen. Additionally, the buttons show numbers that should be kept in mind.

Participants showed a similar error rate in memorization for fingerpointing (m = 2.82, SD = 1.66) and highlighting (m = 2.73, SD = 1.62) (Fig. 4 (right)). The results show that short sequences of numbers (<4) can be kept in mind better when they are shown to the user by highlighting. Thus there is evidence that there is no significant tendency toward a positive effect of fingerpointing for memorizing short sequences (<4). This finding is in-line with [21] who found that pointing positively effects the memorization of a final position, but the cognitive absorption based on the hand movement negatively impacts the memorization. However, for sequences >4 there seems to be a tendency in favor of the fingerpointing, since the amount of errors was then equal or smaller than the errors for the highlighting condition (Fig. 4 (right)). Thus, our findings do not clearly support the results from [4], who also described the supportive character of gestures for memorization of mathematical contexts, which were, however, not related to the learning of sequences.

4.3 Evaluation of Questionnaires

The results from the NASA TLX questionnaire showed a slightly higher perceived task load when using the HoloLens (m = 2.55, SD = 1.20) compared to the highlighting procedure (m = 2.08, SD = 0.88), which comes from the fact that in particular for smaller grid sizes the finger pointing was less easy to detect. This might also be due to the limited field of view, which made pointing gestures to appear more suddenly by just seeing the finger but not the forearm in the peripheral field of view. Although the users also had to wear the HoloLens during the highlighting study, this effect of a limited field of view did not occur, since the HoloLens was switched off. The reasons from the above might also be responsible for the outcomes of the SUS questionnaire, where the HoloLens was rated worse (m = 83.41, SD = 11.74) compared to highlighting (m = 91.59, SD = 7.44). Further analysis with a one-tailed paired samples t-test reveals a statistically significant difference between the usability of the HoloLens compared to highlighting, $t(10) = 2.3$, $p = 0.023$. With regard to cognitive absorption, HoloLens had a higher level (m = 5.04, SD = 0.76, $t(10) = 1.5$) than highlighting (m = 4.67, SD = 0.86). This result is intuitive considering the fact that MR is more immersive than a laptop screen and thus increases users' cognitive absorption.

5 Summary and Outlook

We showed that there is no need to switch from explaining hand gestures to artificial symbols when precise pointing on a specific object is needed. As long as the target object is considerably larger (e.g., $> 29 \times 29$ mm), regular pointing gestures with the finger can be unequivocally detected. Thus, using an MR overlay of hand gestures for explaining, e.g., the operation of machines is sufficient and no further need for pointing arrows is required. This allows for a more natural and intuitive explanation and better understanding of complex contexts. The paper further showed that the intuitive nature of hand gestures also allows for better memorization of longer instructional sequences, since human beings have better access to hand gestures than to other more artificial means of pointing.

Future work will focus on reducing the pointing error further which also stems from sources such as gaze point calibration of the HoloLens when recording or replaying the pointing gestures. Moreover, we will also investigate whether a combination of pointing gestures and highlighting will improve both—the pointing accuracy as well as the memorization of longer instructional sequences. Another study will also focus on the memorization of sequences other than numbers, such as objects, colors, or shapes, to which users might have a more intuitive access.

References

1. Agarwal R, Karahanna E (2000) Time flies when you're having fun: cognitive absorption and beliefs about information technology usage. MIS Q 24(4):665–694
2. Aigner R, Wigdor D, Benko H, Haller M, Lindbauer D, Ion A, Zhao S, Koh J (2012) Understanding mid-air hand gestures: a study of human preferences in usage of gesture types for hci. Microsoft Res TechReport MSR-TR-2012-111 2:30 (2012)
3. Alaçam S (2014) The many functions of hand gestures while communicating spatial ideas-an empirical case study. In: 18th Conference of the iberoamerican society of digital graphics, online, CUMINCAD pp 106–109
4. Aldugom M, Fenn K, Cook SW (2020) Gesture during math instruction specifically benefits learners with high visuospatial working memory capacity. Cogn Res Principles Implications 5(1):1–12
5. Brooke J (1996) Sus: A "quick and dirty" usability. Usability Eval Ind 189(3):189–194
6. Chen S, Chen M, Kunz A, Yantac AE, Bergmark M, Sundin A, Fjeld M (2013) Semarbeta: mobile sketch-gesture-video remote support for car drivers. Christian SABAH 4th augmented human international conference. USA, ACM, New York, pp 69–76
7. Gao Q, Sun Q (2015) Examining the usability of touch screen gestures for older and younger adults. Human Factors 57(5):835–863
8. Hart SG, Staveland LE (1988) Development of nasa-tlx (task load index): results of empirical and theoretical research. Adv Psychol 52:139–183
9. Hoover M, Miller J, Gilbert S, Winer E (2020) Measuring the performance impact of using the microsoft hololens 1 to provide guided assembly work instructions. J Comput Inf Sci Eng 20(6):061001
10. Huang W, Alem L, Tecchia F, Duh HBL (2018) Augmented 3d hands: a gesture-based mixed reality system for distributed collaboration. J Mult User Interfaces 12(2):77–89

11. Kervegant C, Raymond F, Graeff D, Castet J (2017) Touch hologram in mid-air. ACM SIGGRAPH 2017 emerging technologies. NY, USA, ACM, New York, pp 1–2
12. Komine S, Nakanishi M (2013) Optimization of gui on touchscreen smartphones based on physiological evaluation–feasibility of small button size and spacing for graphical objects. In: International conference on human interface and the management of information, Springer, pp 80–88
13. Laviola E, Gattullo M, Manghisi VM, Fiorentino M, Uva AE (2022) Minimal AR: visual asset optimization for the authoring of augmented reality work instructions in manufacturing. Int J Adv Manuf Technol 119(3):1769–1784
14. Leitão R, Silva PA (2012) Target and spacing sizes for smartphone user interfaces for older adults: design patterns based on an evaluation with users. In: 19th conference on pattern languages of programs
15. Lin J, Harris-Adamson C, Rempel D (2019) The design of hand gestures for selecting virtual objects. Int J Human-Comput Int 35(18):1729–1735
16. Liu X, Thomas GW, Cook SW (2018) The effect of pointing on spatial working memory in a 3d virtual environment. Appl Cogn Psychol 32(3):383–389
17. Oyama E, Tokoi K, Suzuki R, Nakamura S, Shiroma N, Watanabe N, Agah A, Okada H, Omori T (2021) Augmented reality and mixed reality behavior navigation system for telexistence remote assistance. Adv Robot 35(20):1223–1241
18. Palmarini R, Erkoyuncu JA, Roy R, Torabmostaedi H (2018) A systematic review of augmented reality applications in maintenance. Robot Comput-Interact Manuf 49:215–228
19. Park YS, Han SH, Park J, Cho Y (2008) Touch key design for target selection on a mobile phone. In: Proceedings of the 10th international conference on human computer interaction with mobile devices and services, pp 423–426
20. Prilla M, Janßen M, Kunzendorff T (2019) How to interact with augmented reality head mounted devices in care work? a study comparing handheld touch (hands-on) and gesture (hands-free) interaction. AIS Trans Human-Comput Interact 11(3):157–178
21. Rossi-Arnaud C, Longobardi E, Spataro P (2017) Pointing movements both impair and improve visuospatial working memory depending on serial position. Mem Cogn 45(6):903–915
22. Schedlbauer M (2007) Effects of key size and spacing on the completion time and accuracy of input tasks on soft keypads using trackball and touch input. Proc Human Factors Ergonom Soc Ann Meet 51(5):429–433
23. Scurati GW, Gattullo M, Fiorentino M, Ferrise F, Bordegoni M, Uva AE (2018) Converting maintenance actions into standard symbols for augmented reality applications in industry 4.0. Comput Ind 98:68–79
24. Teo T, Lee GA, Billinghurst M, Adcock M (2018) Hand gestures and visual annotation in live 360 panorama-based mixed reality remote collaboration. In: Proceedings of the 30th Australian conference on computer-human interaction, pp 406–410
25. Thomas P, David W (1992) Augmented reality: an application of heads-up display technology to manual manufacturing processes. Hawaii international conference on system sciences, vol 2. NY, USA, ACM, New York, pp 659–669
26. Tsang S, Chan A, Chen K (2013) A study on touch screen numeric keypads: effects of key size and key layout. In: International Multi-conference of engineers and computer scientists, vol 324
27. Yusof C, Halim N, Nor'a M, Ismail A (2020) Finger-ray interaction using real hand in handheld augmented reality interface. In: IOP conference series: materials science and engineering, vol 979. UK, IOP Publishing, Bristol, p 012009

Explainable Loan Approval Prediction Using Extreme Gradient Boosting and Local Interpretable Model Agnostic Explanations

S. M. Mizanur Rahman and Md. Golam Rabiul Alam

Abstract Loan approval is an extremely crucial and time-consuming process for a bank. Both customers and bankers can benefit greatly from a loan prediction system since the system helps to reduce time and improve the accuracy of information. However, machine learning models tend to have BlackBox characteristics and bankers cannot understand the internal decision-making process. This study intends to solve this specific issue and examines actual bank loan approval data and adapted a variety of machine learning algorithms for comparative study to select the most appropriate framework for learning information about bank loan approval. The system provided an accuracy of above 95.58% in forecasting with Extreme Gradient Boosting. Moreover, to reveal the key characteristics that determine if a client's loan will be approved or not, Local Interpretable Model Agnostic Explanations have been integrated. This innovation will help bankers to take accurate decisions and also lessen the opacity and fragility of traditional machine learning models.

Keywords Credit risk · Random Forest · Extreme Gradient Boosting · Decision Tree · AdaBoost · Logistic Regression · Loan approval prediction

S. M. Mizanur Rahman (✉)
Bangladesh University of Professionals, Dhaka 1216, Bangladesh
e-mail: mizan78@gmail.com
URL: https://bup.edu.bd/

Md. Golam Rabiul Alam
BRAC University, Dhaka 1212, Bangladesh
e-mail: rabiul.alam@bracu.ac.bd
URL: https://bracu.ac.bd/

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_63

1 Introduction

The disbursement of loans to the clients is one of the main business activities of banks. For the majority of the banks, loan allocation is the main source of income. The major portion of a bank's profit are mostly derived from the income from the loans given to the clients. Nevertheless, banks only provide such loans to applicants who might manage to repay them in order to turn a profit, lowering the likelihood of defaulting. Thus, risk management and determining who is creditworthy are numerous challenges for the banks. The ability to calculate a client's risk level based on variables like profession, gender, relationship status, income range, creditworthiness, etc. is a crucial step which banks take before offering credit to clients. A low credit score or a shallow credit profile, insufficient income, erratic employment, or a mismatch between the use of loan and the lender's loan purpose requirements are all reasons why loan application may be declined. After a lengthy process of verification and validation, many banks nowadays accept loans, although there is no guarantee that the chosen applicant is the most worthy candidate among all applicants. Frauds are a current issue that many banks struggle with as well.

In recent years, academics have investigated the use of several machine learning techniques in credit evaluation/loan approval, including neural networks, support vector machines, decision trees, and integration algorithms. Few of the study have managed to reach state-of-the-art accuracy. Despite reaching higher accuracy financial institutions do not rely on these developed model and continue to approve the loans manually. The main reason behind this is the black box characteristics of the machine learning models. Financial institution do not feel comfortable leaving their most crucial decision on these model without knowing how a model is making its prediction. One of the goal of this article is to solve this particular issue.

This study's contribution may be summarized as follows:

- A comprehensive study between various traditional machine learning models is presented that can classify weather a loan should be approved or not.
- We first ever introduce the interpretability in predictive loan approval decision using Local Interpretable Model Agnostic Explanations.

In this article's Sect. 2, a brief summary of prior studies on bank loan approval is provided. Section 3, which is broken down into three subsections, then provides a brief review of our methodology, models, and strategies. Section 3.1 explains the system model, while 3.2 describes data collection and preparation techniques and 3.3 reveals our model specification. Confusion matrices and performance measurements and findings along with LIME explanation are shown in Sect. 4. Finally, with Sect. 5, we conclude our study.

Table 1 Related works

Refs.	Algorithm	Dataset	Accuracy	Interpretability
Turkson et al. [5]	Random forest	I-Cheng Yeh from UCI	81(±2)	No
Turkson et al. [5]	AdaBoost	I-Cheng Yeh from UCI	81(±2)	No
Eweoya et al. [7]	Decision tree	Private source	75.9	No
Wang et al. [6]	Random forest	Private source	94.57	No
Wang et al. [6]	Decision tree	Private source	92.11	No
This study	**XGBoost**	Li et al. [8]	**95.58**	**Yes**

2 Related Works

Numerous research on analyzing financial data and categorizing loan status have been carried out using various methodologies.

Li et al.'s research employed a weighted-selected attribute bagging approach, and explored the use of customer characteristics to evaluate credit risk [1]. Researchers experimentally compared their results to other state-of-the-art approaches employing two credit databases, and they reported exceptional performance in terms of prediction accuracy and stability. Moro et al.'s proposal to forecast the success or failure of telemarketing for a Portuguese retail bank also includes a data mining technique [2]. They used a variety of data mining models to analyze bank telemarketing data and concluded that neural network data mining was the most effective way. By contrasting four distinct hybrid machine learning algorithms, C. Tsai and M. Chen employed a hybrid machine learning methodology to analyze credit ratings [3]. They conducted experiments to demonstrate that the hybrid "classification Plus classification" model, which combines logistic regression with neural networks, yields the greatest prediction accuracy while also maximizing profit (Table 1).

In order to compare and select the machine learning algorithms that are most appropriate for learning bank credit data, Turkson et al. looked at real bank credit data and run many machine learning algorithms on it [5]. Over 80% of predictions made by the algorithms were accurate. Additionally, out of a total of 23 variables, the most crucial features that predict whether a client would default or not on paying their credit the next month are extracted. The most crucial features were then applied to a few chosen machine learning algorithms, and their predicted accuracy was compared to that of other algorithms that made use of all 23 features. The results indicate no discernible difference, demonstrating that these characteristics can reliably assess clients' credit eligibility.

Wang et al.'s study primarily compares the results of five well-known classifiers used in machine learning for credit scoring: Naive Bayesian Model, Logistic Regression Analysis, Random Forest, Decision Tree, and K-Nearest Neighbor Classifier [6]. It is bold to claim that one classifier is superior to another when each has strengths and weaknesses of their own. However, this experiment's findings show that Random Forest outperforms the competition in terms of precision, recall, area under the curve (AUC), and accuracy.

Our methodology is complementary but distinct in many ways from that used in this discussion. In order to forecast the trustworthiness of a bank's credit data, we used a variety of machine learning techniques. Additionally, unlike the studies mentioned above, our study focuses on the interpretability of the proposed model along with the classification of the loan status.

3 Methodology

From Sect. 3," which is divided into three subsections, we get a clear picture of our proposed model. Section 3.1 addresses our system concept. Section 3.2 covers acquisition and description, and part Sect. 3.3 covers our utilized machine learning models.

3.1 Proposed System Model

We initially began by gathering information on bank related data from numerous sites. Finally, after finalizing our dataset, we started with data preprocessing. Since the acquired data appeared to be fairly noisy, we have considered a variety of preprocessing techniques to clear the data. We have performed a comprehensive EDA in order to find the relation between the features. Additionally, there were a lot null values which needed to be handled along with various data types. After using our data cleaning approaches, we applied several models, including Decision Tree Classifier, Random Forest Classifier, XGBoost Classifier, Logistic Regression, and AdaBoost Classifier. We employed ROC curve and confusion metrics to evaluate the effectiveness of our model. We also utilized a range of performance metrics such as accuracy, recall, precision, and F1-scores to identify which model tends to be more effective. Finally, based on our study, we export the best fitted model that is most appropriate. Local Interpretable Model Agnostic Explanations (LIME), an explainable AI framework, is then used to explain the exported model. It can reveal obscure knowledge that lies behind the model's predictions (Fig. 1).

3.2 Data Acquisition and Description

The data collection is based on information from U.S. Small Business Administration. According to SBA Overview and History, US Small Business Administration (2015), the U.S. SBA was established in 1953 with the goal of promoting and aid-

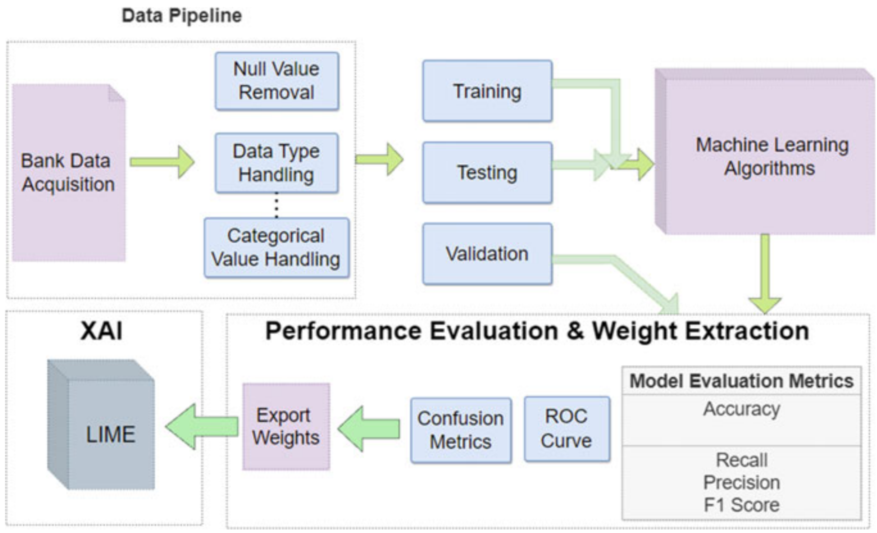


Fig. 1 Proposed system model

ing small businesses in the country's credit market. Therefore, encouraging small business development and growth offers social advantages through generating job opportunities and lowering unemployment. Small businesses have historically been the main driver of job creation in the United States. This dataset is consisted of 27 features. The dataset contains a total of 899,164 records. Figure 5 depicts the distribution of two class. Here, the orange region is represented by good loan and the blue region is depicted as bad loan. The imbalance in the dataset in its entirety is fairly evident.

Figures 2, 3 and 4 represents the analysis performed while preprocessing our dataset. Here Fig. 2 reveals the correlation between the features. Figure 3 represents the number of paid/defaulted loan from 1984 to 2010. We can visualize that the number of paid loan is much greater than the defaulted loan. Finally, Fig. 4 shows the average days till loan disbursement.

3.3 Model Specification

XGBoost A gradient boosting ensemble machine learning algorithm called XGBoost is based on decision trees. It is often used in a multitude of areas because to its high efficiency and accuracy. Using the second order derivative as an approximation, XGBoost attempts to lessen the overall model's error. In all XGBoost scenarios, it has been shown that scaling is due to a number of important frameworks and algorithm improvements, including an unique tree learning algorithm, a logical quantile

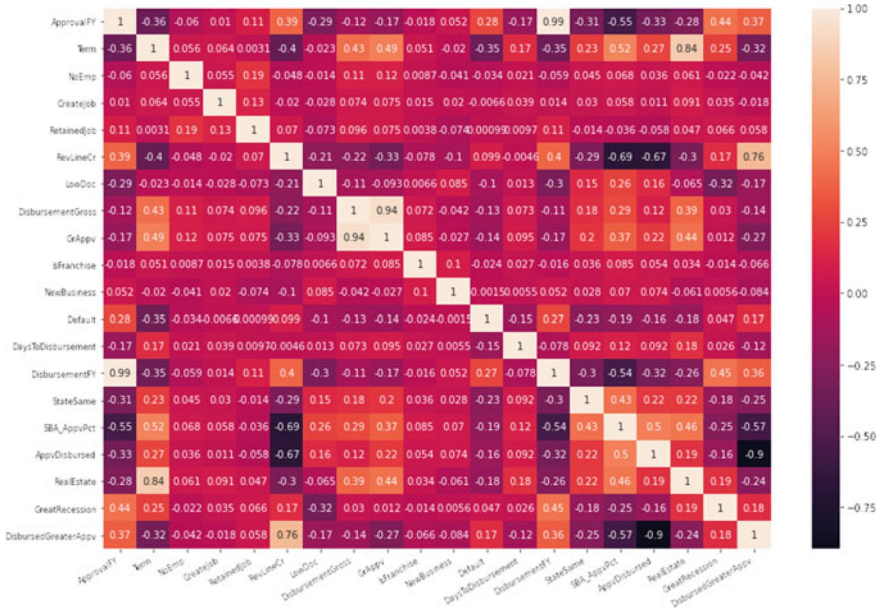


Fig. 2 Data analysis: correlation between features

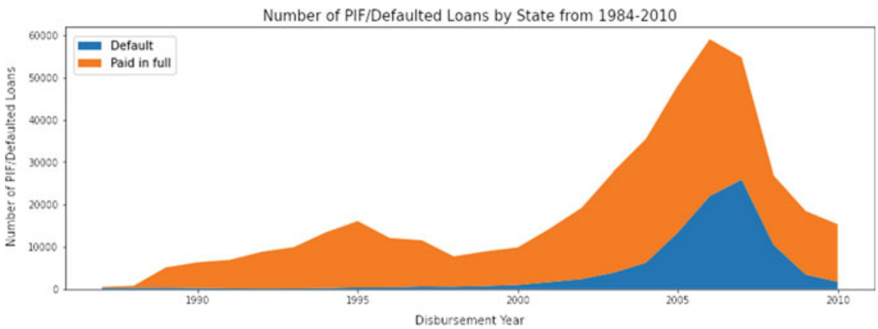


Fig. 3 Data analysis: number of paid/defaulted loan from 1984–2010

sketching technique, and parallel, distributed computing [9]. One of Xtreme Gradient Boosting's (XGBoost) most useful features is the application of parameters that actively change the classifier and improve its accuracy or ease of understanding. Regularization, parallel processing, tree pruning, and the ability to increase learnability are some of XGBoost's best features. First off, XGBoost's optimization functions aid in the learner's increased learnability. The inclusion of regularization techniques like Ridge and Lasso regression, which lessen model bias, is XGBoost's second most useful feature. In essence, regularization is a technique that modifies the predicted coefficients to decrease the variance and sample error. Regularization, also known as

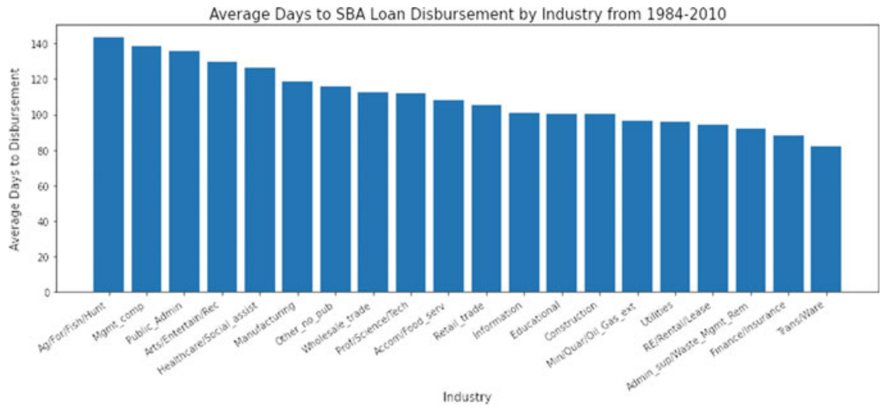
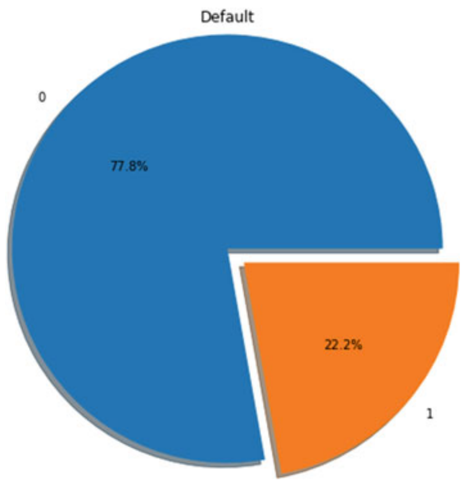


Fig. 4 Data analysis: average days till loan disbursement

Fig. 5 Loan status. Here approved loan is represented by 0 and declined loan are represented by 1



regularized regression, has been presented to achieve generalization, early halting, sparsity manipulation, etc.

Local Interpretable Model Agnostic Explanations The Local Interpretable Model Agnostic Explanations (LIME), which denotes that it is used to explain certain machine learning model predictions [10]. It is one of the most common and well-liked XAI approaches. As a result of manipulating the input data, a synthetic data stream that only contains a small subset of the original properties is produced. LIME is an explainable AI method that works with a variety of classifiers and may be used with text, picture, and tabular data. Equation provides access to the LIME's explanation Eq. 1.

$$\xi(x) = \operatorname{argmin} \quad \zeta(f, g, \pi_x) + \Omega(g) \qquad [10] \tag{1}$$

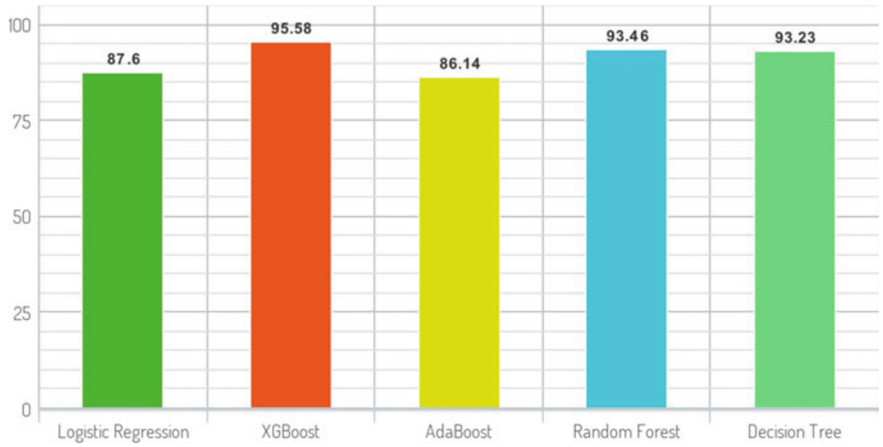


Fig. 6 Performance evaluation: accuracy

4 Performance Metrics and Evaluation

Performance Metrics How precisely a state's predictions reflect the actual world is how accurate it is. It is expressed as a percentage of all forecasts that the framework accurately forecasted. Accuracy becomes more important when true positives and true negatives are more important than false negatives and false positives. In this study, a true positive refers to an instance in which our system correctly finds anomalies, whereas a false positive refers to an instance in which our system incorrectly detects discrepancies.

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{All Predictions}} \tag{2}$$

The method to assess a system's performance is provided by the Eq. 2. In this instance, the total number of predictions is divided by the number of accurate predictions. A model's accuracy is a measure of how frequently, out of all the predictions it made, the model produced the accurate forecast. For instance, a prediction accuracy of 90% means that out of 100 predictions, 90 came true.

Precision can tell the difference between true positives and false positives. When our system predicts that it will find an anomaly but the prediction is unreliable, a false positive will result.

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

The accuracy or correctness of all of a model's positive predictions serves as a measure of its precision. Additionally, it might be used to assess how well the algorithm predicts negative values when the equation is utilized Eq. 3.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{3}$$

Precision and recall both depend on relevance. Recall is the percentage of correct predictions out of all the accurate forecasts that should have been made.

$$\text{Recall} = \frac{\text{True Positive}}{\text{Total Actual Positive}} \tag{4}$$

As par Eq. 4, Recall is defined as the proportion of correct predictions divided by the total number of right predictions it should have made. For example, if a model has a recall value of 60%, it indicates that the model is capable of making 60 right predictions out of the 100 it should have generated.

The F1-score offers a respectable balance of precision and recall even if it is not frequently used as a measure of accuracy, precision, or memory. The F1-score is calculated using a test's accuracy and recall. The F1-score is more important than accuracy if false negatives and false positives have a significant impact, according to the Eq. 5.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

When recall and accuracy are flawless, the F1-score becomes 1. Additionally, this is the highest F1-score that can be obtained. The worst-case scenario, in which neither the accuracy nor recall are present, results in an F1-score of 0.

Performance Evaluation To assess the dataset, we've trained a number of machine learning algorithms. These algorithms produced results that were satisfactory and more accurate. A sarcasm detection model may be simply trained using our supplied dataset, according to the findings of numerous assessment measures, including precision, recall, and F1-scores, which are also employed.

Figure 6 represents our finding after training with various machine learning algorithms. We have utilized five different machine learning algorithms that are Logistic Regression, XGBoost Classifier, AdaBoost Classifier, Random Forest Classifier, and Decision Tree Classifier. Among them with 86.14% AdaBoost Classifier's accuracy was the lowest. With 87.6% Logistic Regression too performed similarly like AdaBoost. The remaining three algorithms managed to acquire more than 93% accuracy whereas XGBoost managed to reach state-of-the-art accuracy with 95.58%.

Table 2 represents the precision, recall, and F1-score of the utilized machine learning models. Here, it is clearly visible that although Random Forest, Decision Tree managed to score close to XGBoost, in terms of their precision, recall, and F1-score XGboost performed more moderately.

Moreover, Fig. 7 depicts the confusion matrix of each of the utilized machine learning models. Here, we can visualize that XGBoost has more True Positive and False Negative value than any other models. Furthermore, Fig. 8 represents the ROC

Fig. 7 Performance evaluation: confusion matrix

Table 2 Performance evaluation

Framework	Precision	Recall	F1-score
Logistic Regression	78.31	60.71	68.40
XGBoost Classifier	91.14	88.71	89.91
AdaBoost Classifier	7504	5660	6453
Random Forest Classifier	88.05	81.07	84.41
Decision Tree Classifier	84.68	85.04	84.86

curve for each of the algorithms. Finally, Fig. 9 illustrates the overall finding of each models. Here accuracy is represented against its precision, recall, F1-score in order to demonstrate the stability of the framework.

LIME Explainability Despite the proposed method's XGBoost classifier's high precision, it is still important for the classification result to be interpretable. Thus we have implemented LIME, which add explainability to our system. Figure 10 displays the four inputs, their predictions, and the LIME prediction. Here, A and D is predicted as 0 or approved loan by our model and LIME identifies the underlying

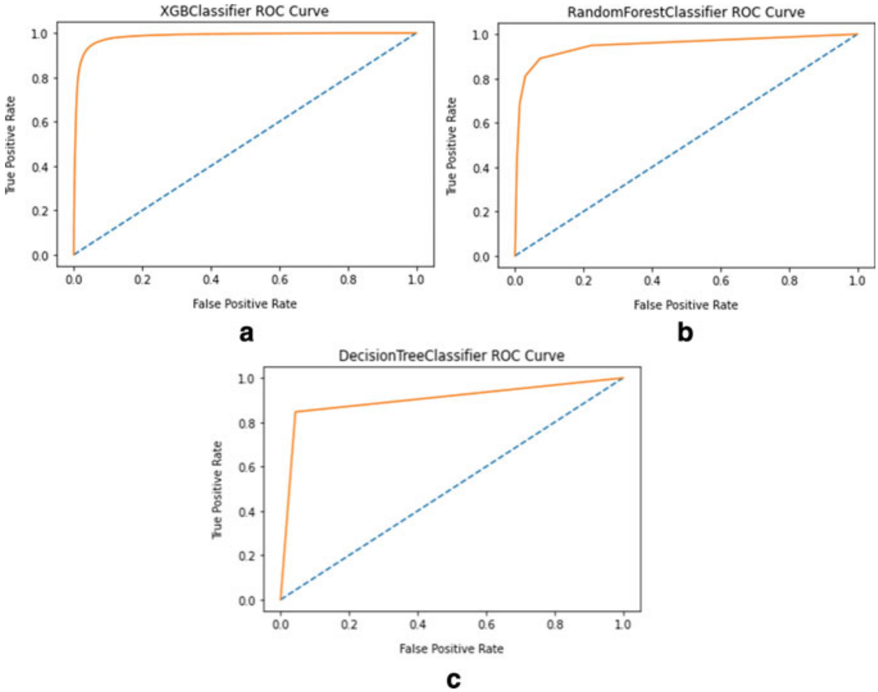


Fig. 8 Performance evaluation: ROC curve of **a** XGBoost classifier, **b** random forest classifier and **c** decision tree classifier

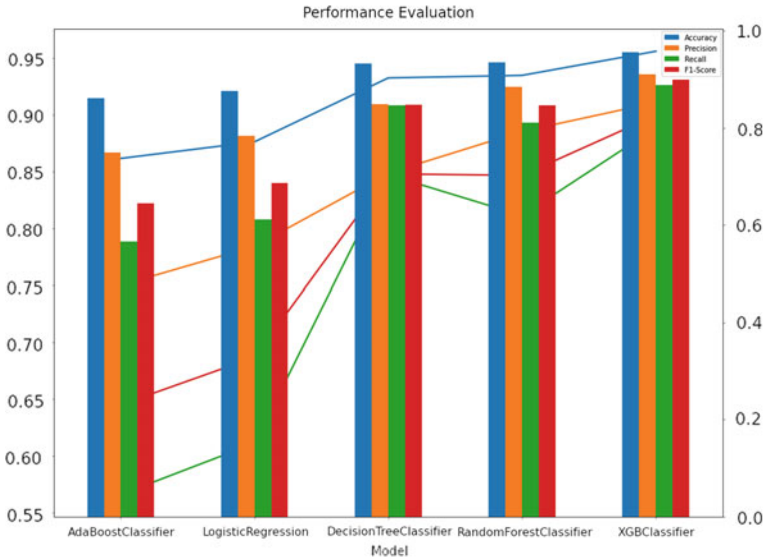


Fig. 9 Performance evaluation: overall metrics comparison

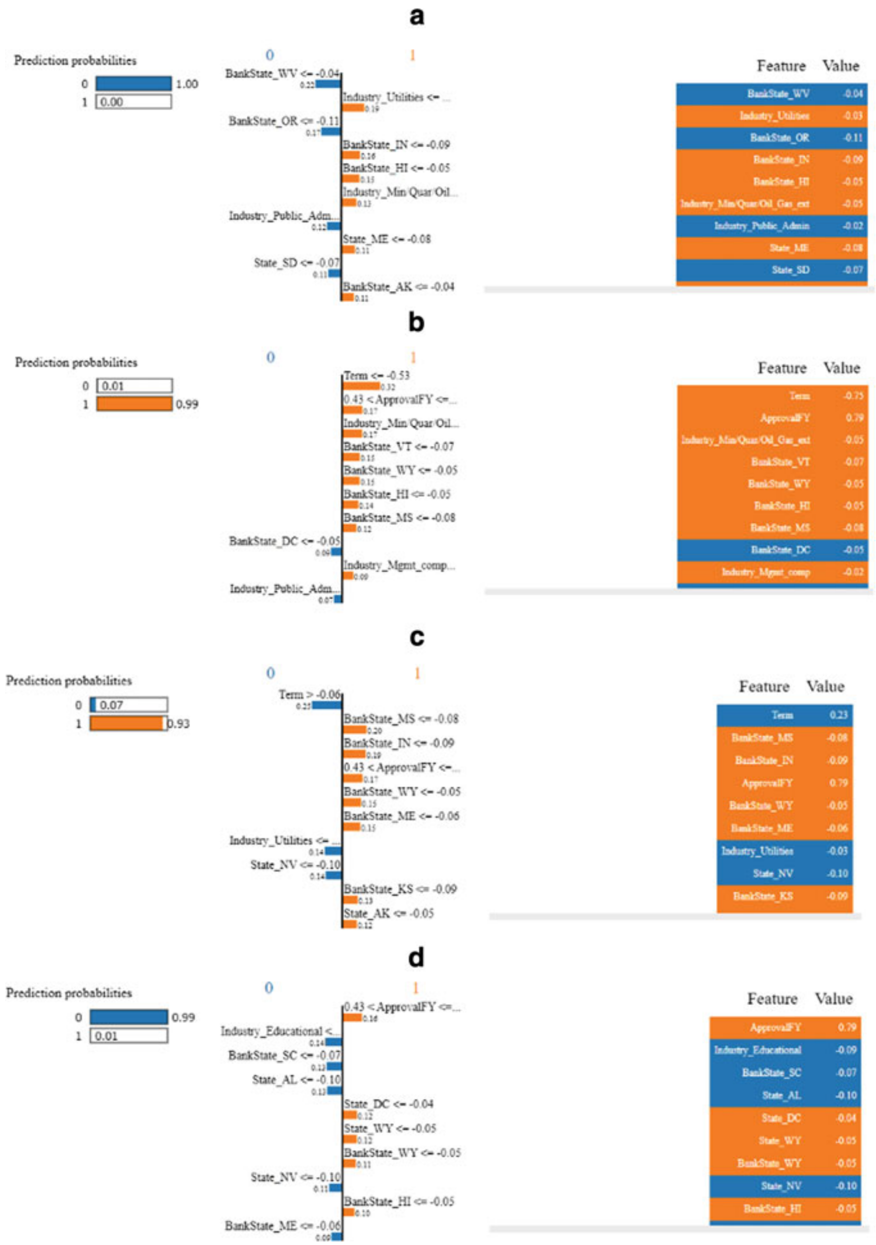


Fig. 10 Local interpretable model agnostic explanations. Here, **a**, **d** interprets the loan approval decision; **b**, **c** interprets the loan declined decision

features behind the prediction. Likewise, B and C are predicted as 1 or declined loan by our model and the influencing features are highlighted as well through LIME.

5 Conclusion

This article proposes a comprehensive machine learning based architectural framework for determining if a loan should be authorized. 5 different machine learning techniques were used on the dataset to identify which algorithms are the most effective ideal for analyzing bank loan datasets. This study reveals that, in addition to the Extreme Gradient Boosting and Random Forest and the other algorithms legitimately outperform each other. According to other performance metrics such as accuracy metrics, precision, recall, and F1-score. These algorithms each attained a certain accuracy rate ranging from 86% to 95%. Additionally, we determined which key factors that affect a person's loan application through Local Interpretable Model Agnostic Explanations. This study has a lot of implications. The model may be used to assist banks in drawing conclusion on approving loans. Additionally, the outcome revealed which algorithms for machine learning performed poorly and are not appropriate for research. Our goal is to create a hybrid machine learning framework that predicts and includes the most crucial characteristics that indicate a customer's loan application worthiness.

References

1. Li J, Wei H, Hao W (2013) Weight-selected attribute bagging for credit scoring. Math Prob Eng
2. Moro S, Cortez P, Rita P (2014) A data-driven approach to predict the success of bank telemarketing. Decis Support Syst 1(62):22–31
3. Tsai CF, Chen ML (2010) Credit rating by hybrid machine learning techniques. Appl Soft Comput 10(2):374–80
4. Tam KY, Kiang MY (1992) Managerial applications of neural networks: the case of bank failure predictions. Manage Sci 38(7):926–47
5. Turkson RE, Baagyere EY, Wenya GE (2016) A machine learning approach for predicting bank credit worthiness. In: Third international conference on artificial intelligence and pattern recognition (AIPR), pp 1–7. https://doi.org/10.1109/ICAIPR.2016.7585216
6. Wang Y, Zhang Y, Lu Y, Yu X (2020) A comparative assessment of credit risk model based on machine learning-a case study of bank loan data. Procedia Comput Sci 1(174):141–9
7. Eweoya IO, Adebiyi AA, Azeta AA, Azeta AE (2019) Fraud prediction in bank loan administration using decision tree. Int J Phys: Conf Ser 1299(1):012037 (IOP Publishing)
8. Li M, Mickel A, Taylor S (2018) "Should this loan be approved or denied?": a large dataset with class assignment guidelines. J Stat Educ 26(1):55–66

9. Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIKDD international conference on knowledge discovery and data mining, 13 Aug 2016, pp 785–794
10. Ribeiro MT, Singh S, Guestrin C (2016) “Why should i trust you?” Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 13 Aug 2016, pp 1135–1144

Evaluating a Synthetic Image Dataset Generated with Stable Diffusion

Andreas Stöckl

Abstract We generate synthetic images with the "Stable Diffusion" image generation model using the Wordnet taxonomy and the definitions of concepts it contains. This synthetic image database can be used as training data for data augmentation in machine learning applications, and it is used to investigate the capabilities of the Stable Diffusion model. Analyzes show that Stable Diffusion can produce correct images for a large number of concepts but also a large variety of different representations. The results show differences depending on the test concepts considered and problems with very specific concepts. These evaluations were performed using a vision transformer model for image classification.

Keywords Image generation · Image classification · Image dataset · Wordnet

1 Introduction

Current models for synthetic image generation cannot only produce very realistic-looking images but also deal with a large number of different objects. In this paper, we use the example of the model "Stable Diffusion" to investigate which objects and types are represented so realistically that a subsequent image classification assigns them correctly. This will give us an estimate of the models potential with regard to realistic representation.

Pre-trained models, such as the one we present, also form the basis for further finetuning to specific objects, as described in [29], and only need a few images of the object. The prerequisite, however, is that the basic model can cope with the desired objects and object classes.

With "Stable Diffusion", we use a current model for image generation to create an artificially generated dataset for training image processing systems. We then evaluate the model using image classification. For the categorization of classes, we use the same approach as ImageNet [5], which uses nouns from Wordnet [15].

A. Stöckl (✉)
University of Applied Sciences Upper Austria, Hagenberg, Austria
e-mail: andreas.stoeckl@fh-hagenberg.at

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_64

For the subset of our dataset that corresponds to the classes in the ImageNet large-scale visual recognition challenge (ILSVRC) [30], we test with an actual image classification method to see how well our synthetic images can be classified. For this, we use the Pytorch implementation of the vision transformer vit_h_14 model from [6] which has a top 1 accuracy of 88.55% and a top 5 accuracy of 98.69% on the real ImageNet data.

This synthetic data is also a good way to improve the diversity of data in a supervised learning setting. They help to get more data without the time-consuming labelling process. Synthetic data can also be seen as the logical extension of data augmentation (e.g. [9, 18, 33]), which is standard in image processing. Seib et al. [32] give an overview of different approaches to enriching real data with synthetic data.

2 Related Work

The "Stable Diffusion" model [26] we use is the latest representative of diffusion models for image generation. The basis of these models was the work of [34] which was improved in [8, 35]. Important and well-known other implementations are "Google Imagen" [31], "GLIDE" [16], "ERNIE-ViLG" [38], "DALL-E" [23], "DALL-E" [4], and "DALL-E 2" [22]. Examples of other image generators are "Midjourney" (https://www.midjourney.com/) and "Google Parti" [37].

Borji [3] investigates how well images generated by DALL-E 2 and Midjourney perform in object recognition and visual question answering (VQA) tasks and compares the results with those on real ImageNet images. The results for the synthetic images are significantly worse, and the authors conclude that "deep models struggle to understand the generated content".

In [2], Stable Diffusion, Midjourney, and DALL-E 2 are examined to see how well they perform in generating faces. They find that Stable Diffusion generates better faces than the other systems. Furthermore, a dataset containing images of faces is provided.

For the training of object recognition methods [7, 21, 28] and segmentation [13, 27], the use of different synthetic image data has been common for some time. Here, the use of synthetic image generators, as described above, offers a variety of further possibilities.

A project that provides access to synthetic image data generated with Stable Diffusion is "Lexica" (Fig. 1—https://lexica.art/). It is a search engine that returns results for a term from over 10 million images. However, the entire database cannot be downloaded here, and there is no categorization.

A large database of 2 million images, which can also be downloaded and used as open source, is offered and described in [36]. Besides the images, the dataset "DiffusionDB" also contains the text prompt used to generate each image, as is the case in our collection. Since this database consists of images that were created by many different users and settings in Stable Diffusion, in contrast to ours, these settings are also stored for each image.

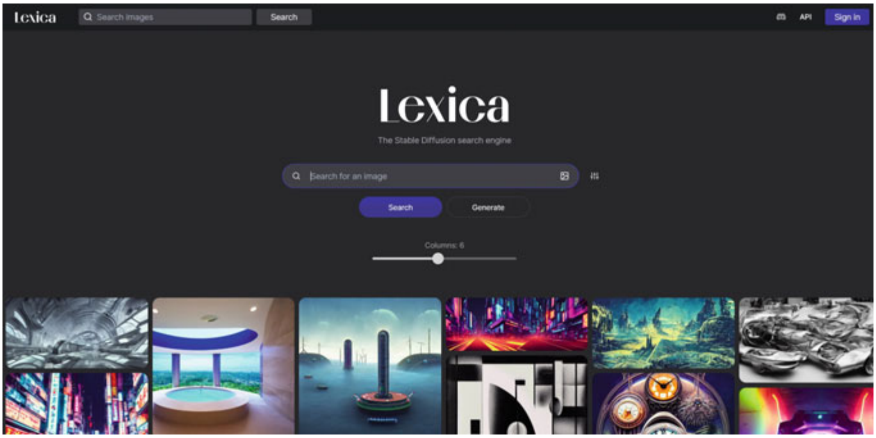


Fig. 1 Lexica.art a synthetic image search

The data collection was created by the authors crawling the discord server of Stable Diffusion and extracting the images including the prompt. Unlike our collection, this does not result in systematic coverage of the wide range of possible concepts, but rather a bias towards the applications that were of interest to the testing community. It also lacks the hierarchy that we have available through the use of Wordnet and use for analysis. The potential applications of "DiffusionDB" that are discussed focus on prompt engineering and explanations and studies of deepfakes.

3 Generation of the Data

As a basis for image generation, we use the "Stable Diffusion" 1.4 model with their Huggingface Diffusers library [19] implementation. This is a model that allows images to be created and modified based on text prompts. It is a latent diffusion model [26] trained on a subset (LAION-Aesthetics) of the LION5B text to image dataset and uses the pretained text encoder CLIP ViT-L/14 [20] to encode the text prompts, as proposed by Imagen [31].

Figure 2 shows an example of an image generated from the text prompt "haflinger horse with short legs standing in water". The example shows that the generator model can represent different concepts with varying attributes and can also combine them in a setting. We now want to create a dataset that contains images of a variety of different concepts in order to evaluate the results.

For text input, we use the information contained in "Wordnet" [15]. Wordnet organizes concepts into so-called "synsets", which correspond to a meaning of one or more words with the same meaning. A word with different meanings can thus belong to several synsets. For example, the word "apple" has the meanings of a fruit and a computer brand, each with a synset for these concepts.

Fig. 2 Image for the text "haflinger horse with short legs standing in water"

In addition to the name, further information is contained for each synset, such as a unique wordnet number and a definition. A directed graph spans between the synsets, which represents the relationships "hypernyms" (a word with a broad meaning constituting a category into which words with more specific meanings fall) and "hyponyms" (a word of more specific meaning than a general or superordinate term applicable to it) and thus makes the hierarchical relationships derivable by being able to output superordinate terms and subordinate terms of a concept.

Starting from the Wordnet synset "object.n.01", we build a list of 26,204 synsets of nouns by recursively calling the "hyponyms". For each of these nouns, we use the description of the synsets in Wordnet for the text prompts of the image generator. For each synset, 10 images are generated and stored under the name of the synset with the number appended. This results in a total of 262,040 images for our dataset. The default settings for Stable Diffusion are 512×512 pixels, 50 inference steps, Guidance Scale 7.5, and PLMS sampling [11]. On an RTX3090 GPU, it took about 6 seconds to create an image. This resulted in a total time of more than 18 days for the creation of the whole dataset.

An example of such a prompt is: (synset for dogs) "a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times occurs in many breeds"

Together with the 10 images per synset, a text file is saved that contains the name of the synset (e.g. "dog.n.01") and the wordnet number (e.g. "n12345678") in addition to the prompt used. The dataset can be downloaded from Kaggle https://www.kaggle.com/datasets/astoeckl/stable-diffusion-wordnet-dataset.

4 Results

First, let us look at some examples of generated images. Figure 3 shows the images generated for the term "Coucal", Fig. 4 for the term "Soccer Ball". It shows on the one hand that very realistic photos were created, and on the other hand a large variety in the representation.

Figure 5 shows the attempt with the abstract term "frame buffer". Here, the model naturally finds it difficult to generate suitable images.

Fig. 3 Generated images for "Coucal"

Fig. 4 Generated images for "Soccer Ball"

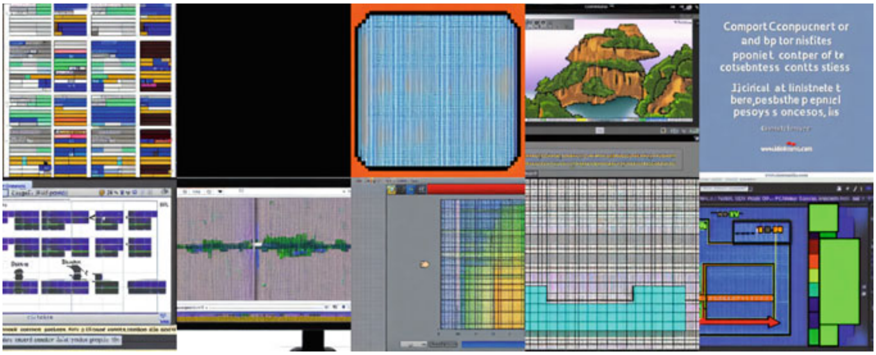


Fig. 5 Generated images for “frame buffer”

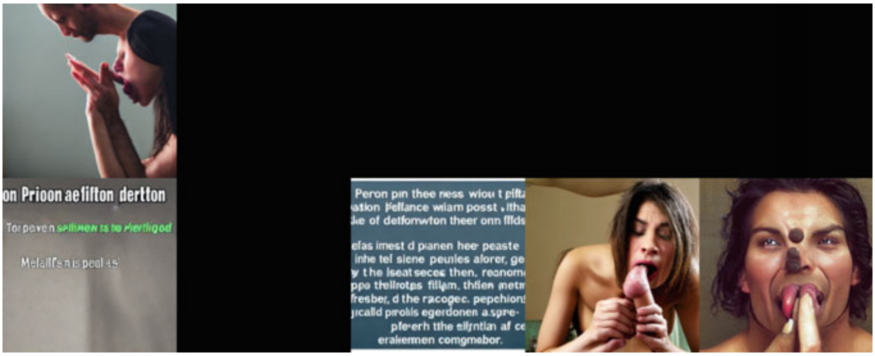


Fig. 6 Generated images for “Cocksucker”

4.1 NSFW Filter

Stable Diffusion has a safety filter that is supposed to prevent the generation of explicit images. Unfortunately, the functionality of the filter is obfuscated and poorly documented. In [24], it was found that whilst it aims to prevent sexual content, it ignores violence, gory scenes, and other similarly disturbing content.

Our tests with sexual content have shown that the filter does not work reliably here either (see Fig. 6 for “Cocksucker”). A black image indicates that the filter has suppressed the output.

An example of filters that have triggered incorrectly is shown in Fig. 7 for the term “System Clock”.

We examine which classes trigger the filter and therefore generate black images. In total, 4620 black images were generated. This is a percentage of 1.76% over all images.

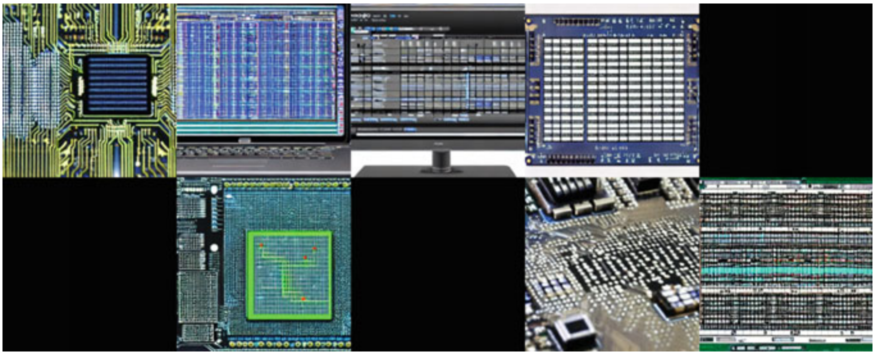


Fig. 7 Generated images for "System Clock"

Fig. 8 Number of correct classified images per class

4.2 Classification with Vision Transformer

We not only to look at and evaluate individual images, but to perform systematic evaluations for a subset of our dataset that is included in the ImageNet large-scale visual recognition challenge (ILSVRC) [30]. That is 861 classes. We use the Pytorch implementation of the vision transformer vit_h_14 model from [6], which has a top 1 accuracy of 88.55% and a top 5 accuracy of 98.69% on the ImageNet data, to verify that the generated images can be correctly classified.

A review of all 8610 images from the considered subset yields a average correct classification of 4.16 images per class (maximum 10) with a average standard deviation of 3.74, across all classes. The histogram in Fig. 8 shows the large spread in the number of correct classifications. The black images generated by the NSFW filter are part of the statistics.

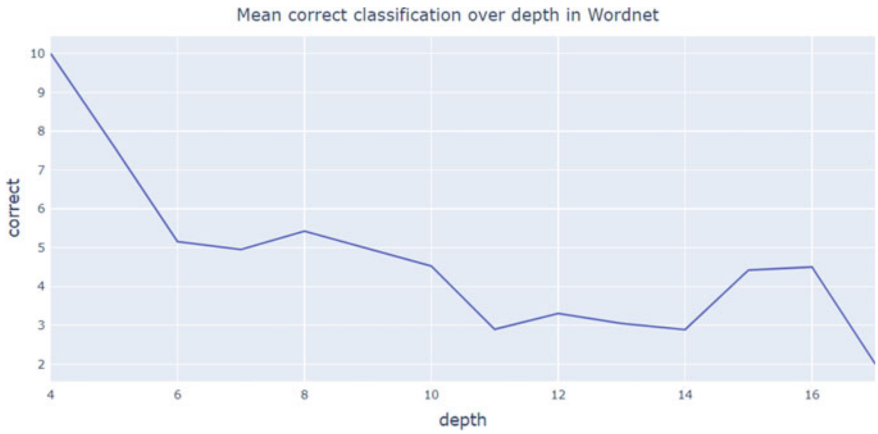


Fig. 9 Number of correct classified images over “depth” in Wordnet

It can be seen that although at least one correctly recognized image was generated for a large part of the classes (73%), only for 14% of the classes all 10 images were recognized again. This also reflects the observation made at the beginning of Sect. 4 on the basis of some examples that the generated images of a class differ strongly. This complicates the task for the classification procedure.

In the Wordnet Taxonomy, there is a “depth” parameter that specifies how many steps you have to descend from the base class to get to the given class. It is therefore a measure of how specific a class is. We now investigate the dependence of the classification rate on the depth.

Looking at the mean recognition rate as a function of depth, the picture of Fig. 9 emerges, indicating a slightly decreasing recognition rate with increasing depth. Generated images of more specific concepts are thus somewhat more difficult to classify correctly.

Let us now consider the recognition rate of some groups of objects. Using the hierarchy of Wordnet, the associated classes were combined for some groups of concepts, and the average recognition rate was determined in each case. Table 1 shows the results.

Remarkable are the good recognition rates for buildings. Figures 10 and 11 show the images for the term “Restaurant” and “Monastery”, where 5 each were correctly classified. Figure 12 shows the images for “Greenhouse” all 10 of which were correctly recognized.

The “Animal” superclass shows below-average classification rates. If we look a little closer at this group, we see that for 162 animal classes, no image was recognized at all. And that the average depth in the Wordnet hierarchy for animals with 12.2 is higher than the overall average of the test classes with 10.4. The test set thus not only contains a particularly large number of classes 376 for animals but also particularly specific ones. These may have made detection more difficult.

Table 1 Recognition rate of different object classes

Group	Number of classes	Mean	Std.
Vehicle	61	4.95	3.79
Animal	376	2.72	3.35
Machine	14	4.14	4.26
Fruit	8	4.5	3.55
Building	11	7.18	2.64
Tool	13	3.85	3.46
Mean	861	4.16	3,74

Fig. 10 Generated images for "Restaurant"

Fig. 11 Generated images for "Monastery"

Looking at individual specific examples, such as Fig. 13 (showing an example of the term "black footed ferret") and Fig. 14 (showing an example of the term "leafhopper"), "Stable Diffusion" obviously reveals significant deficiencies in animal science.

To further overview the results for groups of concepts, we consider some visualizations in the next section.

Fig. 12 Generated Images for “Greenhouse”

Fig. 13 Generated images for “black footed ferret”

Fig. 14 Generated images for “leafhopper”

4.3 Visualization with Word Embeddings

To create a “map” of the terms that shows which of the images generated by Stable Diffusion are correctly recognized by the vision transformer model, and how good the recognition rate is in each case, we plot the terms by semantic meaning in 2D and colour each by subgroup. The size of a circle indicates the number of correctly classified images. To determine the positions on this map, we use word embeddings [14] for the names of the order classes. We use the “fast text model” [1] that was pretrained on Google News and Wikipedia data, since it is trained on the subword level, and the embeddings of the terms can be composed from these. This also avoids that our tested terms are not present in the vocabulary. Using “Gensim” [25], the model was loaded, and the 300 dimensional vectors were extracted.

For the two-dimensional representation, a dimension reduction is necessary for which we used PCA [10] and TSNE [12]. Scikit Learn [17] was used for the computation. For labels consisting of more than one word, the embeddings of the individual words are added to obtain an embedding of the object.

Figure 15, for example, shows the “map” coloured for the superclass “animals” and projected by PCA. Here, too, the many small red dots corresponding to classes that were not correctly recognized are noticeable, as described in the previous section. Figure 16 shows the superclass “building” projected using TSNE. The different classes are not very well represented here by embedding in a common region. The good classification rate of “buildings” is shown by the relatively large red circles in the figure.

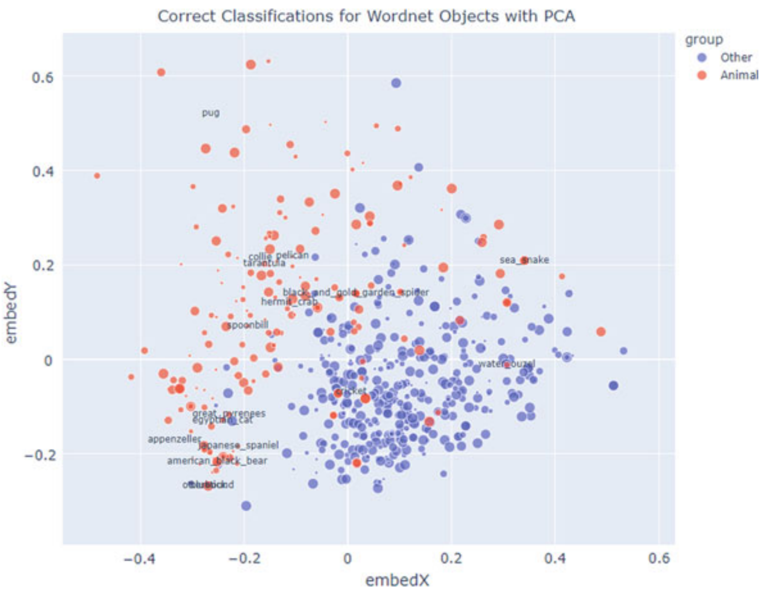


Fig. 15 Map of correct classified images for animals with PCA

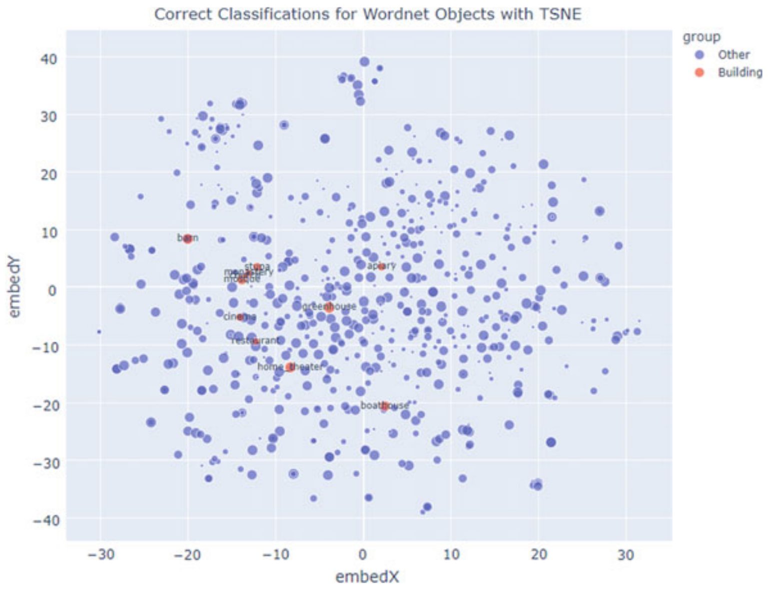


Fig. 16 Map of correct classified images for buildings with TSNE

5 Summary and Future Work

Using the Wordnet taxonomy, it was possible to automatically generate synthetic images for a wide spread set of concepts by using the definitions of the concepts as a prompt for Stable Diffusion.

This dataset can now be used as a basis for various image processing applications that use it for data augmentation. It would be interesting to see if image classification or object detection methods can benefit from this data augmentation. It would also be interesting to train an image classification model like vision transformer on our synthetic dataset or an even larger one and test it on real data.

A second aspect for which the dataset can be useful is to better analyze and understand the generation system "Stable Diffusion". Our first analyzes show that Stable Diffusion generated at least 1 correct image for 10 trials for a wide range of concepts (73%). So for a large part of the world, there is information in the system. On the other hand, very different images are generated for one concept, which is useful for a generative system to be creative, but this also decreases the recognition rate.

We have also seen that different groups of concepts were "understood" differently, as could be seen for example with very specific animal species. There is plenty of room for further investigation and evaluation here.

Finally, it has been shown that the system's filter for unwanted content does not work reliably.

References

1. Bojanowski P, Grave E, Joulin A, Mikolov T (2016) Enriching word vectors with subword information. arXiv:1607.04606
2. Borji A (2022) Generated faces in the wild: Quantitative comparison of stable diffusion, midjourney and DALL-E 2. arXiv:2210.00586
3. Borji A (2022) How good are deep models in understanding\\the generated images? arXiv:2208.10760
4. Dayma B, Patil S, Cuenca P, Saifullah K, Abraham T, Le Khac P, Melas L, Ghosh R (2021) Dall.e mini. https://doi.org/10.5281/zenodo.5146400, https://github.com/borisdayma/dalle-mini
5. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) ImageNet: a large-scale hierarchical image database. In: CVPR09
6. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S et al (2020) An image is worth 16x16 words: transformers for image recognition at scale. arXiv:2010.11929
7. Hinterstoisser S, Lepetit V, Wohlhart P, Konolige K (2018) On pre-trained image features and synthetic images for deep learning. In: Proceedings of the European conference on computer vision (ECCV) workshops, pp 0–0
8. Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. Adv Neural Inf Process Syst 33:6840–6851
9. Inoue H (2018) Data augmentation by pairing samples for images classification. arXiv:1801.02929
10. Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. Philos Trans Roy Soc A: Math Phys Eng Sci 374(2065):20150,202
11. Liu L, Ren Y, Lin Z, Zhao Z (2022) Pseudo numerical methods for diffusion models on manifolds. arXiv:2202.09778
12. Van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. J Mach Learn Res 9(11)
13. McCormac J, Handa A, Leutenegger S, Davison AJ (2017) Scenenet rgb-d: can 5 m synthetic images beat generic imagenet pre-training on indoor segmentation? In: Proceedings of the IEEE international conference on computer vision (ICCV)
14. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. arXiv:1301.3781 (2013)
15. Miller GA (1995) Wordnet: a lexical database for English. Commun ACM 38(11):39–41
16. Nichol A, Dhariwal P, Ramesh A, Shyam P, Mishkin P, McGrew B, Sutskever I, Chen M (2021) Glide: towards photorealistic image generation and editing with text-guided diffusion models. arXiv:2112.10741
17. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V (2011) Scikit-learn: machine learning in python. J Mach Learn Res **12**, 2825–2830
18. Perez L, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning. arXiv:1712.04621
19. Platen von P, Patil S, Lozhkov A, Cuenca P, Lambert N, Rasul K, Davaadorj M, Wolf T (2022) Diffusers: state-of-the-art diffusion models. https://github.com/huggingface/diffusers
20. Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, et al (2021) Learning transferable visual models from natural language supervision. In: International conference on machine learning, PMLR, pp 8748–8763
21. Rajpura PS, Bojinov H, Hegde RS (2017) Object detection using deep CNNs trained on synthetic images. arXiv:1706.06782
22. Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M (2022) Hierarchical text-conditional image generation with clip latents. arXiv:2204.06125
23. Ramesh A, Pavlov M, Goh G, Gray S, Voss C, Radford A, Chen M, Sutskever I (2021) Zero-shot text-to-image generation. In: International conference on machine learning, pp 8821–8831. PMLR

24. Rando J, Paleka D, Lindner D, Heim L, Tramèr F (2022) Red-teaming the stable diffusion safety filter. arXiv:2210.04610
25. Rehurek R, Sojka P (2011) Gensim–python framework for vector space modelling, vol 3, issue 2. NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic
26. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B (2022) High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 10684–10695
27. Ros G, Sellart L, Materzynska J, Vazquez D, Lopez AM (2016) The synthia dataset: a large collection of synthetic images for semantic segmentation of urban scenes. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3234–3243
28. Rozantsev A, Lepetit V, Fua P (2015) On rendering synthetic images for training an object detector. Comput Vis Image Understand 137:24–37
29. Ruiz N, Li Y, Jampani V, Pritch Y, Rubinstein M, Aberman K (2022) Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation. arXiv:2208.12242
30. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) Imagenet large scale visual recognition challenge. Int J Comput Vis 115(3):211–252
31. Saharia C, Chan W, Saxena S, Li L, Whang J, Denton E, Ghasemipour SKS, Ayan BK, Mahdavi SS, Lopes RG et al (2022) Photorealistic text-to-image diffusion models with deep language understanding. arXiv:2205.11487
32. Seib V, Lange B, Wirtz S (2020) Mixing real and synthetic data to enhance neural network training—A review of current approaches. arXiv:2007.08781
33. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1):1–48
34. Sohl-Dickstein J, Weiss E, Maheswaranathan N, Ganguli S (2015) Deep unsupervised learning using nonequilibrium thermodynamics. In: International conference on machine learning. PMLR, pp 2256–2265
35. Song Y, Ermon S (2019) Generative modeling by estimating gradients of the data distribution. In: Advances in neural information processing systems, vol 32
36. Wang ZJ, Montoya E, Munechika D, Yang H, Hoover B, Chau DH (2022) DiffusionDB: a large-scale prompt gallery dataset for text-to-image generative models. arXiv:2210.14896
37. Yu J, Xu Y, Koh JY, Luong T, Baid G, Wang Z, Vasudevan V, Ku A, Yang Y, Ayan BK, et al (2022) Scaling autoregressive models for content-rich text-to-image generation. arXiv:2206.10789
38. Zhang H, Yin W, Fang Y, Li L, Duan B, Wu Z, Sun Y, Tian H, Wu H, Wang H (2021) ERNIE-ViLG: unified generative pre-training for bidirectional vision-language generation. arXiv:2112.15283

Can Short Video Ads Evoke Empathy?

Hasrini Sari, Yusri Mahbub Firdaus, and Budhi Prihartono

Abstract Short videos have become an essential tool in digital marketing to increase sales and performance. This study investigates the effectiveness of short video advertisements that evoke empathic emotions. Three factors in the video are considered: point-of-view, location, and audio. The study object is a short video posted by a café. Data were collected using an experimental design method. This study adopts a pre-test and post-test measurement instrument to measure the viewers' empathetic emotions. The pre-test measures participants' trait empathy, and the post-test measures state empathy. The gap between these two kinds of empathy shows the effect of the video. The story is about a man reading a book in a cafe who is feeling annoyed by the loud laughter of a girl sitting nearby. This study uses eight videos as the stimulus. One is the existing video as the control stimulus. Besides the questionnaire, this study uses EEG and an eye tracker to measure the participants' brain waves and visual attention. Twenty-two males and seventeen females participated in this within-subject experiment. The result shows significant differences between indicators of empathic emotion before and after watching the video. Based on the robust regression method, only two factors significantly influence empathy: audio and point-of-view. However, these two factors can only explain the 8.35% variance of state empathy. Further investigation to explore the influence of gender using a t-test shows no significant differences between the two groups.

Keywords Short video ads · Empathy · Neuromarketing

H. Sari (✉) · Y. M. Firdaus · B. Prihartono
Faculty of Industrial Technology Institut Teknologi Bandung (ITB), Bandung, Indonesia
e-mail: hasrini@itb.ac.id

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_65

1 Introduction

Video is a communication medium to distribute synchronized sound and images [8]. Block [3] explains that what we see are moving images when we watch TV, movies, or videos. If the image being viewed is a movie or video, the viewer can feel the emotional effect after watching the movie or video.

Short videos have become an essential tool in digital marketing to increase sales and performance. Short video advertising is gaining popularity especially in social media. Addo et al. [2] state that short video advertising has a significant direct relationship with sales. Ge et al. [11] investigate user-generated short videos in social media. The result shows that music and female vividness significantly affect product sales.

However, many videos are intended not only for sales but also to evoke certain emotions in the viewers. Former studies investigate the effect of video games [13] and movie trailers [10] on emotion. Choe et al. [6] argue that the same video scene evokes different emotions in different individuals. This study investigates factors that can evoke empathy in viewers after watching a short video. The result can help marketers in creating more effective short video ads.

2 Literature Review

2.1 Empathy

Empathy is defined as a person's emotional response based on that person's understanding of another person's emotional stimulus or emotional state [9]. Shen [21] explained that empathy is divided into two: trait empathy and state empathy. Trait empathy is the initial concept of empathy which views empathy as a person's fundamental nature. State empathy is the concept of empathy that arises because of certain conditions, including information. In simple terms, state empathy is the empathy that is generated when someone is faced with a condition. Shen [21] measures the state of empathy after someone processes information from a message. Shen [21] divides empathy into three dimensions: cognitive empathy, affective empathy, and association empathy. Cognitive empathy refers to the condition when a person takes a particular perspective. This perspective-taking process involves recognizing, understanding, and adopting another's point-of-view. Affective empathy is the activation of affective responses to the emotions of others that involves the process of understanding the feelings of others. In comparison, association empathy is the stage where the audience can project themselves into the message or act as if they are part of it. Shen [21] develops a measurement instrument for each of the three dimensions of empathy.

2.2 Information Delivered in a Video

Block [3] explains that three factors influence the audience of a video: story, sound, and visual. A story is a series of events to be conveyed. Sound is a series of dialogue, sound effects, or music used to support the story being told. Meanwhile, visual is a visual description of the story that can be seen. Block [3] adds three aspects that affect the visual of an image: story, viewpoint, and location. The point-of-view describes the emotions of each object in the image. Verleur et al. [23] explained that the audience would see the character's emotions well when the camera is closer to the character's face (close-up). In comparison, location is the place/setting of the story that is described.

According to Ginea and Thomas [12], two factors influence the delivery of information in a video: audio and picture. The experiment in this study was carried out by manipulating the image displacement per second (fps) of the existing video. Based on their research, the audience still obtained the information through the video's sound even though the fps was changed.

2.3 Brain Waves

Nerves in the brain communicate with each other through waves of electrical impulses. There are five brain wave types: Gamma, Beta, Alpha, Theta, and Delta. Each of these waves has a different frequency and different function. Delta wave (0.5–4 Hz) is generated when someone sleeps. Theta (3–8 Hz) exists in deep relaxation and meditation conditions. Alpha (8–12 Hz) exists in relaxation and passive attention conditions. Beta (12–27 Hz) is generated when someone is in an active mind condition, and Gamma (>27 Hz) is when someone concentrates on thinking about something [1]. Ismail et al. [14] found a relationship between human emotions and brain wave activation. Four types of emotions, namely (anger, happiness, emotional surprise, and sadness), are investigated. The result shows that each emotion triggers the production of different types of brain waves. The emotion of anger drives the Theta wave produced. Sadness drives two brain waves, namely Delta and Theta. Happy emotion triggers the generation of the Alpha wave. In contrast, emotional surprise triggers almost every type of brain wave.

3 Methods

This study examines the effectiveness of a short video post on Instagram. The video's objective is to evoke the state empathy of the viewers. This study uses the story as the control variable by referring to the three aspects of a visual image put forward by

Block [3]. Two different point-of-views, two different locations, and the existence of audio are tested.

Therefore, three hypotheses were formulated as follows:

H1: The location in the video has a positive relationship with the state empathy.
H2: The point-of-view in the video has a positive relationship with the state empathy.
H3: The audio in the video has a positive relationship with the state empathy.

The object is a video ad from Café X posted on its Instagram. According to management, the short video has not generated empathy because visitors have not shown the desired behavior. This study focuses on associative empathy because it is a dimension of empathy that connects perception and action. However, because association empathy is a stage after cognitive empathy and affective empathy, in this study, cognitive and affective empathies were still measured by limiting the number of questions. The measurement instrument used is a questionnaire using a 5-point Likert scale which refers to Shen [21]. Participants filled out the questionnaire after seeing the stimulus. While before the experiment, the participants also filled in a questionnaire developed by Corte et al. [7]. Trait empathy reflects empathy which is a person's nature.

Data collection used a muse band and eye tracker. A muse headband is a device used to measure brain waves. The relevant brain wave for measuring empathy is the Theta wave. According to Ismail et al. [14], Theta waves come from the right part of the brain that reflects the emotion of anger. The muse has seven calibrated sensors—two on the front of the head, two on the back of the ear, and three as reference sensors—that detect and measure brain activity.

Eye tracker was used to measure eye movement to find out the participants' areas of interest [4]. There are two types of eye trackers based on how they are used: wearable eye trackers and remote eye trackers. In this study, the type of eye tracker used is a remote eye tracker called the 3GP eye tracker. The device is placed in front of the participant and connected to a computer screen that displays the stimulus.

This study also uses a muse headband in the data collection process. A muse head band is a device used to measure brain waves. The relevant brain wave for measuring empathy is the Theta wave. According to Ismail et al. [14], theta waves come from the right part of the brain that reflects the emotion of anger. This study uses electroencephalogram (EEG) to measure the wave. According to Lal and Craig [15], EEG can provide information about changes in a person's brain condition by measuring brain waves. In addition, Neumann and Westbury [17] stated that EEG has electrodes that can record the activity of neurons in our brain by attaching them to the scalp. This study uses muse and an EEG that measures the brain's work based on the waves it receives. The muse has seven calibrated sensors—two on the front of the head, two on the back of the ear, and three as reference sensors—that detect and measure brain activity.

Data collection uses an experimental design method. The three basic principles of experimental design are [16]: randomization, replication, and blocking. Randomization is the form of randomizing the test sequence to meet certain probability distribution assumptions and reduce systematic errors. Replication is the repetition of a test treatment to allow the estimation of an experimental error (for statistical inference) and increase research precision by reducing random error. Blocking is an activity to eliminate variability due to factors, not of concern to research (nuisance factors).

In this study, there are two factors and two levels for each factor. Levels are denoted by 0 and 1. The factors and levels can be seen in Table 1.

Data were collected using an eye tracker to test the hypotheses of location and point-of-view. Meanwhile, to test the audio hypothesis, using data collected from the muse band.

The type of experiment used is a 2^k factorial design, so the number of required stimuli is $2^3 = 8$. The control variables in this study were the actor/actress, the story, and the duration of the video. The story depicted in this video is that the main character, named Hardianto, was focusing on reading a book, while the supporting character, named Fasya, was sitting nearby and playing with her cell phone and laughing out loud without paying attention to the people around her. Hardianto, who was reading a book, was annoyed and looked at Fasya with annoyance. As for the duration, this video is 8 s long. Examples of the stimuli can be seen in Fig. 1.

The stimulus display method used in this study is within-subject so that participants can compare the stimulus received in the experiment. This study uses a partial counterbalancing approach in determining the sequence of the stimuli to minimize the carry effect.

Table 1 Factors and level

Factors	Level	References
Location	0: outside of the café; 1: 1: inside the café	Block [3]
Point-of-view	0: close-up of the main character; 1: close-up of the supporting character	Block [3]; Verleur et al. [23]
Audio	0: without audio; 1: with audio	Ghinea and Thomas [12]

a b

Fig. 1 Examples of stimuli, **a** existing video (outside, the supporting character, no audio); **b** stimulus (inside, main character, audio)

At last, the experiment procedure and protocol were made as guidance during the experiment. The procedure contains steps of experimenting as follows: (a) The participant enters the experiment room; (b) the participant reads the experiment protocol; (c) the participant has explained the experiment rules; (d) the participant signed the statement of participation; (e) the participant fills in the trait empathy questionnaire; (f) the participant is explained about the scenario of the experiment; (g) The participant wears the muse band and calibrated; (h) eye tracker calibration; (i) stimulus displayed; and (j) the participant fills in the state empathy questionnaire. The protocol describes the experiment's goal, the experiment rules, the statement of willingness to participate, and the scenario of the experiment.

The sample of this study refers to the target market of the café, which are aged between 17–23 years old, students, have more than 2.5 million rupiahs monthly spending, and follow the Instagram of the café. The minimum number of participants for the experiment using an eye tracker is 30 [19], and EEG is 15 [24]. Therefore, this study sets 30 as the minimum number of participants. The data collection location was in the Laboratory of Innovation and Enterprise System Design, Institut Teknologi Bandung. Before the data collection, a pilot test was conducted, and the result was used to refine the experiment procedure and protocol.

4 Data Collection

The data collection process resulted in participant profiles as follows: 56.41% of participants were male, the age of participants ranged from 17 to 23 years, and all were students. In addition, all participants follow the Instagram of Cafe X and have a monthly expenditure of less than IDR 2500.000. The characteristics of these participants are following the target market of the café.

Pre- and post-test questionnaire validity was measured using Pearson correlation, and the result showed that all the indicators have a significant value of less than 0.05. The questionnaire reliability test showed that Cronbach's Alpha values were 0.710 and 0.905. The questionnaires are reliable because according to Malhotra and Birks [16], the measurement instrument is reliable if it has a Cronbach's Alpha value greater than 0.6.

The first step of data processing was data normalization using a min–max method to data collected from the eye tracker and muse band. Normalization was done for the eye tracker because there were two data types: attention duration and frequency. For data from the muse band, normalization was done to investigate changes in the temporal brain of the participants during the stimulus exposure. Furthermore, the gap value between trait and state was also calculated to indicate the level of empathy changes.

The next step was normality data testing using the Kolmogorov–Smirnov test. The results of the data normality test showed that there were data with a significance value of less than 0.05, namely location and empathy. Therefore, it can be said that

Table 2 Output of robust regression analysis

Variable	Coefficient	Std. error	Probability
Audio (A)	13.482	2.584	0.0000
Location (L)	– 0.079	1.154	0.9453
Point-of-view (PV)	– 1.632	0.612	0.0077
Constanta	17.065	1.500	0.0000

the data held are not normally distributed. Therefore, the regression test used is robust regression. The result of the regression analysis can be seen in Table 2.

Table 2 shows two significant variables: audio and point-of-view. Therefore, hypotheses 2 and 3 are accepted. Therefore, the regression equation is:

$$\text{State empathy} = 17.065 + 13.482A - 0.632\text{PV} \tag{1}$$

The adjusted R^2 of the equation is 0.0835 or 8.35%. This value shows that the point-of-view and audio factors can explain the empathy variance of 8.35%. In contrast, other variants of 91.65% are explained by other variables that have not been considered in this study.

The third step is comparing participants' empathy levels before and after seeing the stimulus. The hypothesis was tested using paired t-test twice, for data collected from the questionnaire and muse band. From the point-of-view, data were compared between the video using the supporting and the main character. The hypothesis is as follows: H0: State empathy of the participants is the same for both types of point-of-view.

H1: State empathy of the participants is significantly different for different types of point-of-view.

The result showed that H0 is accepted (p-value 0.119 for the questionnaire and 0.311 for the muse band). Therefore, the differences are not significant.

For audio, the hypothesis is as follows:

H0: State empathy of the participants is the same, with or without audio.

H1: State empathy of the participants is significantly different for video with and without audio.

The result showed that H0 is rejected (p-value 0.00 for the questionnaire and 0.022 for the muse band). Therefore, the stimulus with audio differs significantly from the stimulus without audio, where the mean value for audio is high.

5 Results and Discussion

This study uses two tools: an eye tracker and a muse headband. Eye tracker collects visual data and is used for testing the influence of point-of-view and location on empathy. Muse band collects the Theta wave and is used to test audio's influence on empathy. The weakness of using these tools is that they are susceptible to movement; therefore, calibration should be done several times during the experiment.

Regression analysis shows two factors significantly influence empathy: audio and point-of-view. However, paired t-test to investigate the state of empathy between the stimulus of the main character and the supporting character's point-of-view shows no significant differences. Therefore, the stimulus designed using the main character's point-of-view cannot evoke a significantly different level of empathy relative to the existing video (control). Spencer et al. [22] also finds no significant effect of point-of-view of a step-by-step video to the performance of students with disability.

Data collected from the eye tracker show that the Area of Interest (AOI) of participants on the main character is higher for all stimuli with a supporting character point-of-view. The negative coefficient of point-of-view in the regression equation may indicate that participants give more attention to the main character when the video shoot close-up of the supporting character. The participants' curiosity may cause this condition on the main character's face. At the beginning of the video, the main character's face is not visible because of his back to the camera. Further investigation to explore the effect of participants' gender on response using unpair t-test analysis shows no significant differences (p-value > 0.05). This in line with the previous study from Matern et al. [18] shows that gender does not significantly influence selective attention of young adult on a video game.

This study investigates three independent variables that resulted in a low value of adjusted R^2 (8.35%). Further research can involve other variables, namely lighting, color, and duration, to improve the quality of the research model. Brown [5] explains that lighting and colors used in videos can be used to direct viewers to specific emotions. In comparison, in their research, Verleur et al. [23] used a 50s video to find out the emotions felt by someone. Another thing that can be the focus of future research is how long the emotions caused by watching videos will last. In their study, Palmer et al. [20] found a significant effect of different video durations on the confidence–accuracy relationship in the context of eyewitness identification decisions.

6 Conclusion

This study investigates the factors of a short video published on Instagram to evoke the emotion of empathy. Three factors in the video are considered, namely point-of-view, location, and audio. EEG and an eye tracker are used to measure the participants' brain waves and visual attention. The result shows that audio and point-of-view are significant in evoking empathy. However, the variance explained by these two factors is relatively small.

Further investigation shows that the gap in empathy level before and after watching the video is not significant for stimuli with a different point-of-view. In conclusion, audio in a short video is essential to evoke viewers' emotions. Further study can be done to investigate the influence of lighting, color, and duration. It is also interesting to study the effect of curiosity.

Acknowledgements PPMI 2022 ITB.

References

1. Abhang PA, Gawali BW, Mehrotra SC (2016) Introduction to EEG- and speech-based emotion recognition. Elsevier Inc., London. https://doi.org/10.1016/B978-0-12-804490-2.00002-6
2. Addo PC, Akpatsa SK, Nukpe P, Ohemeng AA, Kulbo NB (2022) Digital analytics approach to understanding short video advertising in digital marketing. J Market Theory Pract 30(3):405–420. https://doi.org/10.1080/10696679.2022.2056487
3. Block B (2008) The visual story: creating the visual structure of film, TV, and digital media. Elsevier Inc., London
4. Bojko A (2013) Eye tracking the user experience: a practical guide to research. Rosenfield, New York
5. Brown B (2012) Cinematography theory and practice: image making for cinematographers & directors. Elsevier Inc., Waltham
6. Choe W, Chun H, Noh J, Lee S, Zhang B (2013) Estimating multiple evoked emotions from videos. In: Proceedings of the annual meeting of the cognitive science society, vol 35(35). ISSN 1069-7977
7. Corte KD, Buysse A, Verhofstadt LL, Roeyers H, Ponnet K, Davis MH (2007) Measuring emphatic tendencies: reliability and validity of the Dutch version of the interpersonal reactivity index. Psychol Belg 235–260
8. Cubitt S (1993) Videography: video media as art and culture. Macmillan Education, New York
9. Cuff BM, Brown SJ, Taylor L, Howat DJ (2014) Empathy: a review of the concept. Emot Rev 144–153
10. Ellis JG, Lin WS, Lin CY, Chang SF (2014) Predicting evoked emotions in video. In: IEEE international symposium on multimedia, pp 287–294. https://doi.org/10.1109/ISM.2014.69
11. Ge J, Sui Y, Zhou X, Li G (2021) Effect of short video ads on sales through social media: the role of advertisement content generators. Int J Advertising 40(6): 870–896. https://doi.org/10.1080/02650487.2020.1848986
12. Ghinea G, Thomas JP (1998) QoS impact on user perception and understanding of multimedia video clips. ACM Multimedia 49–54
13. Hemenover SH, Bowman ND (2018) Video games, emotion, and emotion regulation: expanding the scope. Ann Int Commun Assoc 42(2):125–143. https://doi.org/10.1080/23808985.2018.1442239
14. Ismail WO, Hanif M, Mohamed SB, Hamzah N, Rizman ZI (2016) Human emotion detection via brain waves study by using electroencephalogram (EEG). Int J Adv Sci Eng Inf Technol 1005–1011
15. Lal SK, Craig A (2005) Reproducibility of the spectral components of the electroencephalogram during driver fatigue. Int J Psychol 137–143
16. Malhotra NK, Birks DF (2007) Marketing research: an applied approach, 3rd edn. Prentice Hall, Inc., London
17. Neumann DL, Westbury HR (2011) The psychophysiological measurement of empathy. Psychol Empathy 1–24
18. Matern MF, Westhuizen A, Mostert SN (2020) The effects of video gaming on visual selective attention. S Afr J Psychol 50(2):183–194. https://doi.org/10.1177/0081246319871391
19. Nielsen J, Pernice K (2009) How to conduct eye tracking studies. Nielsen Norman Group, California

20. Palmer MA, Brewer N, Weber N, Nagesh A (2013) The confidence-accuracy relationship for eyewitness identification decisions: effects of exposure duration, retention interval, and divided attention. J Exp Psychol Appl 19(1):55–71. https://doi.org/10.1037/a0031602
21. Shen L (2010) On a scale of state empathy during message processing. West J Commun 504–524
22. Spencer GP, Mechling LC, Ivey AN (2015) Comparison of three video perspectives when using video prompting by students with moderate intellectual disability. Educ Train Autism Dev Disabil 50(3):330–342
23. Verleur R, Heuvelman A, Verhagen PW (2011) Trigger videos on the Web: impact of audiovisual design. Br J Educ Technol 573–582
24. Wiechert G et al (2016) Identifying users and activities with cognitive signal processing from a wearable headband. In: IEEE 15th international conference on cognitive informatics and cognitive computing (ICCI*CC), pp 129–136. https://doi.org/10.1109/ICCI-CC.2016.7862025

Optimize One Max Problem by PSO and CSA

Mohammed Alhayani, Noora Alallaq, and Muhmmad Al-Khiza'ay

Abstract The optimal solution in mathematical concepts, computer science, and finance is to find the best solution out of all possible solutions. The type of optimization problem is determined by whether the variables are continuous or discrete.This study presents a One Max solution that shifts from the notion of optimization to the notion of the optimal strategy. Based on the time dimension difference, the CSA and PSO algorithms have been proposed as more effective in optimization, since the PSO algorithm is the oldest in the optimization field and CSA is modern. Nevertheless, despite being newly configured, the CSA algorithm has proven its effectiveness. Both algorithms must use values that are generated at random. Each cycle has a predetermined range of values for 100, 500, and 1000 cycles, and the values are calculated using the Sigmoid function. They go through 30 cycles with a number of function evaluations of 100,000. The Sigmoid function, which raises values above 0.5–1, is used to display the results for each range of 30 values. The results showed that the CSA algorithm outperformed PSO by 20% in terms of improvement values for each cycle (100, 500, and 1000). The CSA algorithm was selected as the preferred method for improving the One Max problem because of its efficiency and speed. Moreover, it has less dispersion than the PSO algorithm.

Keywords Optimizations · One Max problem · Sigmoid · CSA · PSO

M. Alhayani (✉)
Al-Noor University College, Ninevah-Mosul, Iraq
e-mail: mohammed.yaseen@alnoor.edu.iq

N. Alallaq · M. Al-Khiza'ay
Al-Kitab University, Kerkuk-Altun Kupri, Iraq
e-mail: Noora.h.mohammed@uoalkitab.edu.iq

M. Al-Khiza'ay
e-mail: Mohammed.kazim@uoalkitab.edu.iq

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_66

1 Introduction

Marketing investigators are interested in achieving peak performance, which it necessarily involves personal well-being, self-determination, and efficiency. The working principle for optimum results is challenging, resulting in an out-of-the-ordinary state of effectiveness. The optimization problem in mathematics, computer science, and economics is to discover the optimal answer from all potential solutions. Optimization problems are classified into two types based on whether the variables are continuous or discrete.

This study covers One Max Problem Optimization that maximizes the number of ones in a feasible solution, and the problem itself is quite simple and widely used in the evolutionary computational community. The instructions lead to solving the One Max problem using Particle Swarm Optimization (PSO) and Crow Search Algorithm (CSA). Each program runs 30 times with the number of function evaluations = 100,000. Take upper bound = 1 and lower bound = −1. The problem is solved for three different dimensions which are $D = 100$, $D = 500$, and $D = 1000$. The two algorithms were compared using the following metrics:

Best, Mean, Median, Worst, and Standard Deviation. Also, make statistical analysis using Wilcoxon, etc.

Particle Swarm Optimization and Crow Search Algorithm were presented for solving continuous optimization problems. On the other hand, the One Max problem is a binary optimization problem, so the solution space must be adapted from the continuous domain to the binary domain. It can be used Sigmoid function for this purpose.

Simple steps in the Algorithms:

Firstly, generate a random number between the lower bound and upper bound. Suppose it is −0.3232. Calculate *Sigmoid* (−0.3232) = 0.419. So, if it is less than 0.5, it becomes 0. The objective value is 0 [1].

After that PSO and CSA generate a new solution using the existing solution (0.3252). Suppose it is −0. 5131. Calculate *Sigmoid* (−0.5131) = 0. 374. Again it is less than 0.5, and it becomes 0. The objective value is 0. After that, the two algorithms generate a new solution using the existing solution (−0.5131). Suppose it is 0. 0856. Calculate *Sigmoid* (0.0856) = 0.521, now it is greater than 0.5 so it becomes 1. The objective value is 1.

The following sections represent this paper: Sect. 1: Introduction; Sect. 2: Literature Review; Sect. 3: Methodology; Sect. 4: Experimental Outcomes; Sect. 5: Discussion; and Sect. 6: Conclusion.

2 Related Work

Business researchers are concerned with optimal performance, which necessitates personal well-being, independence, and optimization. The operating mechanism for optimum performance is complicated, which results in an unusual state of performance. Any system's performance status can advance from one level to another, increasing output, effectiveness, and delivery time [2]. The level of optimal performance affects why performance is at its best. Understanding the complexities of optimal functioning, such as how someone achieves optimal cognitive functioning, is novel, especially in terms of educational and social implementation methods, and it advances our understanding of the relationship between optimization and optimal performance [3].

The CSA algorithm has been used by researchers to address a wide range of issues in numerous fields [4]. In order to address integer optimization and minimax problems, this study suggests a new cuckoo search algorithm that combines the cuckoo search algorithm with the Hill-Climbing approach. The suggested strategy is known as Cuckoo and Hill-Climbing Hybrid Search (CSAHC). The Hill-Climb algorithm is used by CSAHC as an intensification process to speed up the search and overcome the slow convergence of the conventional cuckoo search algorithm after the standard Cuckoo search is applied to the number of iterations. By using 13 criteria, performance validation is determined. According to the results of an experimental simulation, CSAHC works better than regular CSA [5]. Our contribution to this study is to suggest other optimization algorithms like PSO to solve the One Max problem and compare them among other criteria.

PSO, a strategy for optimization that was inspired by social animals' group behavior, was one of the techniques the researchers covered. A swarm of particles may flow across the parameter space that specifies the courses pushed by them as their best performers and those of their neighbors, and this is how the set of potential solutions to an optimization problem is established. The ability of Particle Swarm Optimization to resolve various optimization issues in chemical measurements is demonstrated in this work. Through the offered succinct literature survey as well as many other fields of chemical measurement, optimization can be used. It has been demonstrated to be helpful for signal alignment as a result of its capacity to find the ideal orientation in space according to the projection index or for variable selection [6]. It can use PSO for solving the One Max problem for its high ability to spread and determine values in different places.

Researchers have first introduced the Automatic Propulsion Particle Swarm Optimization (CSA–PSO) technology, which serves both electric companies and their customers in terms of economic and environmental benefits. In this study, the allocation, size, and number of urban planning clusters were optimized based on the goals of minimizing overall costs and energy loss [7]. To calculate the decrease in overall costs and total energy losses, a new reduction ratio formula is applied. It is demonstrated that the CSA–PSO method is superior at resolving the optimal power flow problem with RDGs, compared to recent metaheuristic innovations [8]. It can be said that the contribution that PSO and CSA make in the field of improvement and problem solving is the best in all areas.

3 Methodology

This research study provides a more thorough comparison between the PSO and CSA algorithms in order to look into the various optimization approaches for the One Max problem. Without going against restrictions, optimization works toward (the best or most efficient use) of a particular set of parameters. Cost reduction and maximizing productivity and efficiency are the most typical objectives. One of the primary quantitative methods used in industrial decision-making is optimization. Each algorithm is executed 30 times with job evaluation count = 100,000. The upper bound = 1 and the lower bound = −1. The solution to the problem has three different dimensions which are $D = 100$, $D = 500$, and $D = 1000$. The following criteria were used to compare and evaluate the performance of the two algorithms: Best, Mean, Median, Worst, and Standard Deviation. Also, we have statistical analysis using Wilcoxon. The way it works is generate a random number between the lower bound and the upper bound. According to the Sigmoid function, if the value is less than 0.5, the objective value becomes 0. If it is greater than 0.5, it becomes 1. For execution, 30 values will be collected for each of the three dimensions (100, 500, and 1000) of the two algorithms, and a comparison will be made between their output values according to the criteria mentioned above.

3.1 Particle Swarm Optimization (PSO) Algorithm

A stochastic optimization algorithm, introduced in 1995, is a population-based that is driven by the intelligent collective behavior of some animals such as flocks of birds or flocks of fish and ants, and also it has undergone many improvements. It is a technique for computing that enhances an issue by iteratively attempting to raise the quality of a possible solution [9]. By moving a set of potential solutions—here referred to as particles—in the search space in accordance with a straightforward mathematical formula over the particle's position and velocity, it solves problems. Each particle moves toward the best-known positions in the search space, which is updated when other particles find better positions. This movement is governed by each particle's local most well-known position. It is anticipated that this swarm would promote better solutions [10].

These are the definitions of the corresponding update formulas for PSO [11]:

$$v_{ij}(k+1) = w \cdot v_{ij}(k) + c1r1\big(\text{pbest}_{ij}(k) - x_{ij}(k)\big) + c2r2\big(\text{gbest}_{j}(k) - x_{ij}(k)\big) \quad (1)$$

$$x_{ij}(k+1) = x_{ij}(k) + v_{ij}(k+1), \quad j = 1, 2, \ldots, D \tag{2}$$

The current position is $x_{ij}(k)$ for the jth dimension of the ith particle in the kth iteration and $v_{\mathrm{ij}}(k)$ is the velocity for the ith particle in the jth dimension in the kth iteration; $\mathrm{pbest}_i = (\mathrm{pbest}_{i1}, \mathrm{pbest}_{i2}, \ldots, \mathrm{pbest}_{iD})$ is the best position for the ith particle that has ever been searched. W is the inertia weight, which influences how much the particle maintains its initial velocity, determining the tendency to be optimized globally or locally. $\mathrm{gbest} = (\mathrm{gbest}_1, \mathrm{gbest}_2, \ldots, \mathrm{gbest}_D)$ is the best place for which all particles have ever been searched. This model's purpose is to mimic bird behavior, and it is: Each unique bird is represented as a random point in the Cartesian coordinate system with an initial velocity and position. Run the program again, this time using the "nearest proximity velocity match rule," and set each individual's speed to equal that of its closest neighbor. If the iteration is repeated, all of the points will quickly have the same velocity. Because this model is overly naive and divorced from actual settings, an additional random variable is added to the speed component. In other words, in addition to satisfying "the nearest proximity velocity match," each speed will also have a random variable added to it at each iteration, making the overall simulation resemble the real scenario, as shown in the pseudocode below [12, 13].

```
For each particle
        Set position and velocity at random.
End
t = 1
Do
       For each particle
           Determine the fitness function
           If fitness value > pBest Then
                  Set current fitness value as pBest
       End
       Update particle with best fitness value as gBest
       For each particle
            Calculate new velocity using equation (18)
            Update position using equation (19)
       End
t = t + 1
While (t < maximum iterations)
Post process the result.
```

3.2 *Crow Search Algorithm (CSA)*

It is a new metaheuristic optimization method that mimics the cognitive behavior of crow swarms. Askarzadeh introduced this technique in 2016, and preliminary results have shown its capacity to solve numerous complex engineering optimization issues.

It works by simulating birds storing and collecting surplus food as needed using a newly developed swarm intelligence algorithm [14]. In optimization theory, the crow is the researcher, the surrounding environment is the search space, and randomly storing the position of the food is a possible option. The location with the most food is regarded as the universal optimum solution among all food locations, with the quantity of food as the aim function. It works by duplicating the intelligent behavior file of crows, which has gotten a lot of interest because of advantages like simple implementation, a minimal number of parameters, adaptability, and so on [15].

The definitions of the corresponding formulas for CSA are [16]:

$$x_i^{\text{iter}+1} = x_i^{\text{iter}} + r_i \cdot \text{fl}_i^{\text{iter}}\left(m_j^{\text{iter}} x_i^{\text{iter}}\right) \to x_i^{\text{iter}+1} = \text{a random position.} \tag{3}$$

$$x_i^{\text{iter}+1} = \left\{x_i^{\text{iter}} + r_i \cdot \text{fl}_i^{\text{iter}}\left(m_j^{\text{iter}} x_i^{\text{iter}}\right) r_i \geq \text{AP}^{\text{iter}}, \quad \text{a random position.}\right. \tag{4}$$

r_i is an integer number between 0 and 1, and fl_i $^{\text{iter}}$ denotes the flight length of crow i at iteration iter, where AP_i $^{\text{iter}}$ signifies the awareness probability of crow j at iteration iter [17].

3.3 One Max Problem

One Max problem is a simple optimization problem. The aim of the problem is maximizing the number of ones in a feasible solution x. x can be either 0 or 1 [18]. The formula of the problem is given below.

$$\max f(x) = \sum_{i=1}^{D} x_i, \tag{5}$$

where D is the dimension of the problem. One Max problem is a binary optimization problem so that the solution space must be adapted from continuous domain to binary domain. We can use the Sigmoid function for this purpose. The Sigmoid function [19]: If *Sigmoid* $(x) \geq 0.5$, then it becomes 1, otherwise 0.

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}. \tag{6}$$

4 Experimental Results

The results of the values for the PSO and CSA algorithms were collected, and the collection of values was based on implementation thirty times in three different dimensions (100, 500, and 1000). In each execution cycle, random numbers were generated depending on the mentioned dimension range. The resulting value is confined between the lower bound, which is -1, and the upper bound, which is 1. Depending on the Sigmoid function, the output value is made either 1 or zero, and the values confined within the range are summed for each cycle, as in Table 1.

Comparing the results of the two algorithms in the iteration of 100,000, it was found that the CSA has higher and more accurate values than the PSO algorithm for the three dimensions. We can say that optimization using the CSA algorithm has better results. But, it remains only to measure the results on the evaluation criteria mentioned to confirm the conclusions that have been found (Table 2).

It is clear from the evaluation criteria table that the CSA algorithm outperforms PSO in the (Best, Mean, Median, Worst, Wilcoxon) value and also the standard deviation of the dispersion of values. The lower the standard deviation of a dataset, the closer the data is to the mean and the less scattered [20]. If the standard deviation is a large number, this indicates that the dispersion of the data is high. So, the standard deviation is a number to indicate the degree of dispersal of the members of the data set [21]. It has been concluded that the values of the algorithm CSA are less scattered than the algorithm PSO.

5 Discussion

CSA and PSO produced very different results in optimizing the One Max problem in terms of (Best, Mean, Median, Worst, Wilcoxon values, and standard deviation), also the values of the two algorithms in the three dimensions (100, 500, 1000). It was discovered that the CSA algorithm has better and more precise values for the three dimensions than the PSO algorithm. We can say that using the CSA algorithm for One Max problem optimization produces superior outcomes. Additionally, because the values of the CSA algorithm are more uniform than those of the PSO method, the CSA algorithm's evaluation criteria table is superior than the PSO algorithm (Fig. 1).

6 Conclusion

It was found that CSA has higher and more accurate values than the PSO algorithm for the three dimensions. We can say that optimization with CSA algorithm has better results. In addition, the evaluation criteria table of CSA algorithm is superior to PSO as the values of CSA algorithm are less dispersed than PSO algorithm. The CSA and

Table 1 Methodological results for PSO and CSA algorithm

Run number	$D = 100$		$D = 500$		$D = 1000$	
	PSO	CSA	PSO	CSA	PSO	CSA
1	54	80	241	313	520	601
2	48	81	241	320	557	594
3	50	75	266	314	496	587
4	48	82	255	304	488	595
5	53	80	250	323	511	598
6	60	80	252	324	473	602
7	49	76	239	312	488	594
8	52	81	266	307	482	594
9	49	75	240	311	515	592
10	43	79	268	324	501	596
11	49	83	224	308	496	601
12	56	75	243	310	471	594
13	56	78	244	317	492	594
14	50	75	259	309	490	594
15	48	84	262	304	496	586
16	37	82	237	309	524	585
17	50	75	248	321	485	583
18	51	75	254	311	503	588
19	60	78	250	308	507	586
20	53	79	252	316	472	595
21	44	76	252	321	501	599
22	54	73	263	322	472	590
23	50	76	258	309	486	577
24	55	80	257	316	496	602
25	53	73	254	312	454	596
26	52	78	259	317	492	604
27	47	78	243	316	529	596
28	60	76	239	312	502	593
29	47	80	245	317	495	587
30	53	76	260	316	495	596

PSO algorithms were proposed to improve the One Max problem, and the results proved that the CSA algorithm outperformed the PSO in the improvement values by 20% for each cycle (100, 500, and 1000). We conclude from Table 1 that the values collected by default in the CSA algorithm were more effective and higher accuracy of PSO values, many of which have been neglected because they are weak values in the Sigmoid function. As for the criteria table, it was concluded that all the values in

Table 2 Evaluation criteria

Evaluation criteria	$D = 100$		$D = 500$		$D = 1000$	
	PSO	CSA	PSO	CSA	PSO	CSA
1. Best	60	84	268	324	557	604
2. Mean	51	78	250	314	498	593
3. Median	50	78	225	313	496	594
4. Worst	37	73	224	304	454	583
5. Standard deviation	5.01	2.97	10.32	5.76	22.55	6.28
6. Wilcoxon	Statistic = 0.0, p value = 1.6954815515692352e−06	Statistic = 0.0, p value = 2.4297055330462724e−06	Statistic = 0.0, p value = 1.7224282827430733e−06	Statistic = 0.0, p value = 1.705141514598868e−06	Statistic = 0.0, p value = 1.7180929312456739e−06	Statistic = 0.0, p value = 1.6805458207281375e−06

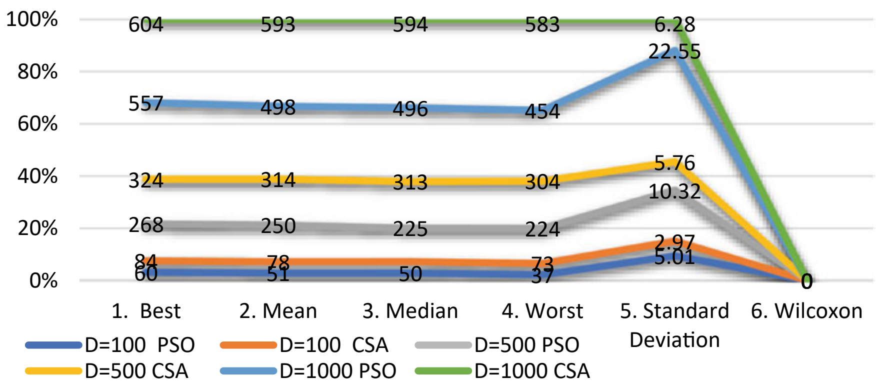


Fig. 1 Evaluation criteria for PSO and CSA

the CSA algorithm were close to the mean values, and this indicates the balance of the algorithm values, as well as the low value of the standard deviation coefficient, from which it was concluded that the amount of dispersion in it is small, unlike the PSO algorithm. Therefore, it is recommended to use the CSA algorithm to improve One Max problem.

References

1. Al-Khiza'ay M et al (2020) Top personalized reviews set selection based on subject aspect modeling. In: International conference on knowledge science, engineering and management, pp 276–287
2. Mujika I et al (2018) An integrated, multifactorial approach to periodization for optimal performance in individual and team sports
3. Phan HP, Ngu BH, Yeung AS (2019) Optimization: in-depth examination and proposition
4. Shehab M, Khader AT, Al-Betar MA (2017) A survey on applications and variants of the cuckoo search algorithm
5. Shehab M et al (2017) Hybridizing cuckoo search algorithm with hill climbing for numerical optimization problems. In: 2017 8th International conference on information technology (ICIT), pp 36–43
6. Marini F, Walczak B (2015) Particle swarm optimization (PSO). A tutorial
7. Askarzadeh A (2016) A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm
8. Farh HM et al (2020) A novel crow search algorithm auto-drive PSO for optimal allocation and sizing of renewable distributed generation
9. Chander A, Chatterjee A, Siarry P (2011) A new social and momentum component adaptive PSO algorithm for image segmentation
10. Wang D, Tan D, Liu L (2018) Particle swarm optimization algorithm: an overview
11. Dai Q, Zhang H, Zhang B (2021) An improved particle swarm optimization based on total variation regularization and projection constraint with applications in ground-penetrating radar inversion: a model simulation study
12. Dai HP, Chen DD, Zheng ZS (2018) Effects of random values for particle swarm optimization algorithm

13. Norouzi H, Bazargan J (2020) Flood routing by linear Muskingum method using two basic floods data using particle swarm optimization (PSO) algorithm
14. Sayed GI, Hassanien AE, Azar AT (2019) Feature selection via a novel chaotic crow search algorithm
15. Hussien AG et al (2020) Crow search algorithm: theory, recent advances, and applications
16. Wu H et al (2020) Finite element model updating using crow search algorithm with Levy flight
17. Li LL et al (2021) Using enhanced crow search algorithm optimization-extreme learning machine model to forecast short-term wind power
18. Frank A, Murota K (2022) A discrete convex min-max formula for box-TDI polyhedra
19. Nantomah K (2019) On some properties of the sigmoid function
20. Divine G et al (2013) A review of analysis and sample size calculation considerations for Wilcoxon tests
21. Lee DK, In J, Lee S (2015) Standard deviation and standard error of the mean

Hospital Information System as a Code Automation and Orchestration

Mohammed Amine Chenouf, Mohammed Aissaoui, and Hafida Zrouri

Abstract The ongoing digital revolution affects individuals and businesses alike. Increasingly, social networks and digital devices are the default means for engaging government, businesses, and civil society, as well as friends and family members. This means that the last best experience that people have anywhere becomes the minimum expectation for the experience they want everywhere, including in the hospitals. This is the domain of digital transformation and its intersection with cloud adoption.

Keywords Hospital information system · Cloud computing · Automation · IAAC · Orchtestration

1 Introduction

Increasing customer expectations and a more competitive business context have placed tremendous pressure on business leaders to change the way they set their strategies and run their organizations. New requirements to incorporate more information and greater interactivity quickly drive-up costs and complexity [1]. This is the domain of digital transformation and its intersection with cloud adoption. Digital transformation incorporates the change associated with the application of digital technology in all aspects of society [2]. Cloud adoption is the way in which businesses implement digital transformation to achieve an end which can be:

- Exceptional user experience
- Accelerated time to market
- Higher service quality

M. A. Chenouf (✉) · M. Aissaoui · H. Zrouri
National School of Applied Sciences, Mohammed Premier University, Oujda, Morocco
e-mail: m.chenouf@ump.ac.ma

M. Aissaoui
e-mail: m.aissaoui@ump.ac.ma

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_67

- Cost flexibility
- Repeatability and flexibility
- Safety, security, and compliance with regulation.

More than that, in this article we will describe the technical architecture for HIS to improve its Information system and costs, the solution is based in microservices architecture using docker images and Infrastructure as a Code (IaaC) [3].

Hospitals needs a suitable infrastructure solution in AWS cloud platform to tackle the complexity of the deployment of the Information system and similar applications in the future. The automation become more and more important for all the deployment, it help the IT administrators to reduce the efforts and time for the build, update and configuration, also it helps to provide automation in order to reduce VM provisioning time [4]. This solution will improve the process of the infrastructure health-check and the resolution of issue, the main goal of this approach is no human action needed [5].

The adoption of our solution enables the hospitals IT administrators to maintain the:

- High Availability
- Scalability
- Automation
- Auditability
- Monitoring

The article consists of four sections that are:

Section one: Deploy a Tool chain allowing autonomy of provisioning through GitOps strategy and infra as code
Section two: Deploy an Eco-system allowing the deployment of microservices applications in a standardized way
Section three: Deploy the AWS fully managed Kubernetes cluster
Section four: this section it's for conclusion and perspective.

2 The GitOps Strategy and Infra as Code

2.1 Introduction

The entire infrastructure will be implemented as a code using Terraform this will track resource changes throughout the infrastructure deployments [6].

In order to enable team collaboration, S3 Backend will be implemented. The state file will be stored remotely in an S3 bucket with Versioning enabled, a Database will be used also for locking to prevent concurrent operations on a single workspace.

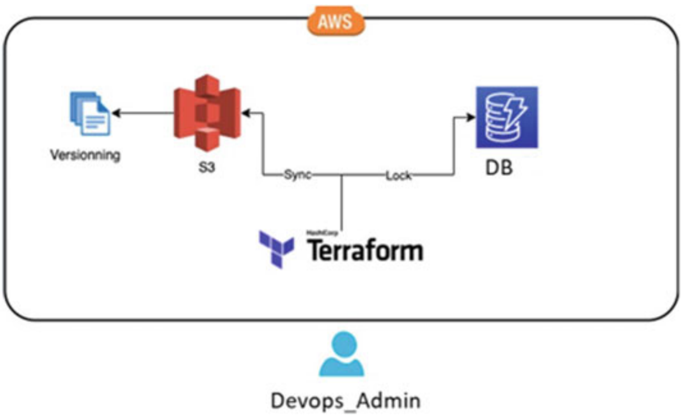


Fig. 1 VMs provisioning work flow on AWS

This new methodogy of build will allow:

- Automate VMs provisioning workflow on AWS
- Replace "Traditional" delivery mode of work by Agile mode of work
- Homogeneity of the VMs
- Agnosticity of the solution to adopt it for other Cloud provider.

Our solution will be deployed on AWS cloud provider, with the combine of Automation platform based on AWS IaaS, Ansible Tower, Terraform, and GitHub SaaS (Fig. 1).

Here are the key architecture parameters which will be taken before the build of a Cloud infrastructure [4]:

- External and Internal Infrastructure Connectivity
- Cloud Accessibility: Platform and then Application
- Core Infrastructure: Workload and Location
- Cloud Management: Monitoring Solution, Retention Policies, Alerting, Scheduling
- Security: Vault, Secret Management, VM Encryption, Traffic filtering
- Backup: Solution, Restoration tests strategy/methodology
- DRP: Solution.

2.2 Security Process

As we work with personal data, we know that the security aspect is the biggest challenge of our topic. to achieve this we decide to have:

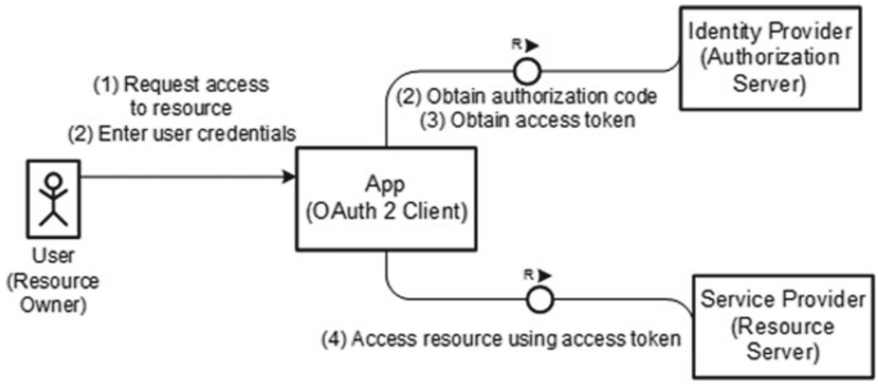


Fig. 2 Authentication process

1. Master Account:

New AWS Organization with new master account in order to activate Security Policies in each client account

2. Security Account:

Security account at the highest level will host audit and security logs from all HIS accounts. Only HIS Security team will have real rights on all logs and will give each account owner read rights on his own logs [7].

- CloudTrail
- CloudWatch Log
- S3 Log repo
- KMS
- Security Appliances.

3. HIS Active directory:

All accounts within HIS OU will follow new guardrails

4. The Authentication Process:

We will use the SSO on AWS to authenticate users. When the user wants to login, a logon API call is received by an Identity Provider which checks if the user is present or not in HIS's AD. It is a process of the same family as OAuth2 Authorization Code Flow [7] (Fig. 2).

If the user is found, the Identity Provider will attribute an STS token valid during a limited period. The user is then redirected to his Corporate SSO to authenticate. If the user is not found, the browser may display an error message.

5. Shared Services with all the VPCs

- Network/Firewalls/Internet Gateway

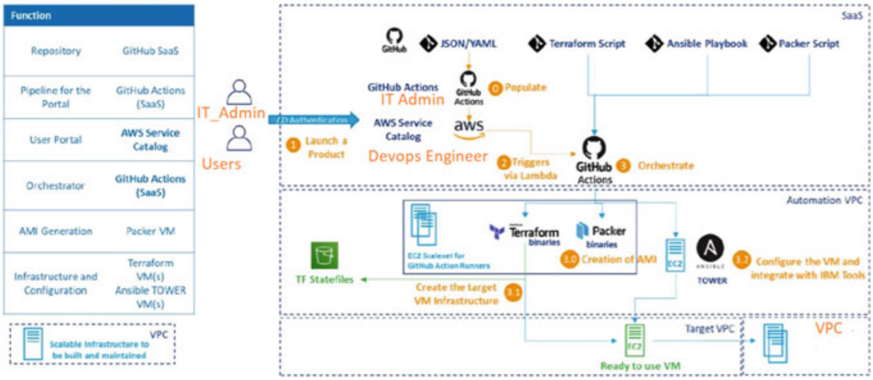


Fig. 3 Technical architecture

- Route 53 Resolver
- DC of HIS
- Bastion
- Future CI/CD Tools.

To simplifies the implementation of the solution, we suggest the below technical architecture.

Figure 3 show the IT admin use and combine the Iaac scripts and devops tools to start the build and launch the configuration of the hospital information system [8].

1. Logging with AWS/AD account
2. AWS attach a fluenbit sidecar automatically to the triggers via lambda
3. Launch the lambda function to configure and build the infrastructures with teraforme, ansible, and git scripts
4. Starting the creation of the Image of the EC2 instances
5. Creating the target EC2 instance
6. Starting the configuration of the Vm with Ansible playbooks
7. Sending an email to the administrator's users with the full information of the infrastructures
8. Starting the work on the target solution

3 The Eco-System and Microservices

The main target of the solution that we suggest is the use of microservices on the HIS to benefit from its advantage and reduce the complexity and the time of the update and the configuration [8]. More than that, it will help to make all the modification on it with less effort and with no downtime of the production. There are many solutions to orchestrate the microservices like: Docker, Kubernetes, Kubernetes ... [9]. For the

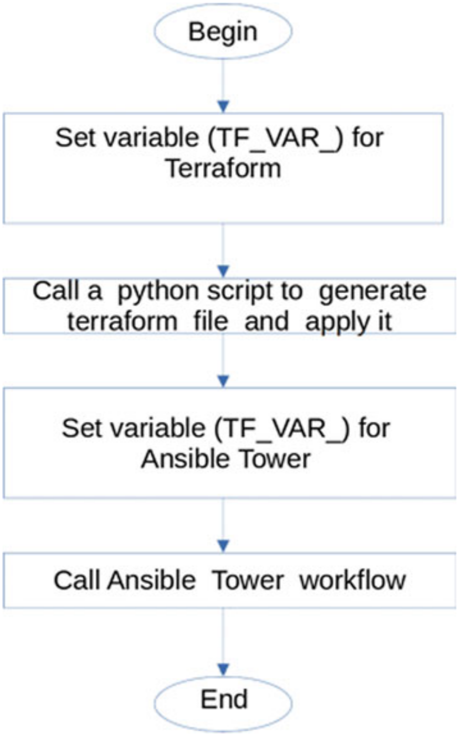


Fig. 4 Lambda flows

solution that we describe, we will use a Kubernetes cluster as an orchestrator solution for our microservices [9].

Lambda functions are also used in the solution, Fig. 4 describe the flows of Lambda function.

After Lambda function the build could be started, Fig. 5 show the steps of the vm provisioning.

3.1 Components of the Eco-System

Our developed terraform scripts, that will be hosted in the CODE repository in GITHUB, should provision and configure the following essential services:

- EKS Cluster
- Pods Profiles
- Proxy
- IAM Policies and Roles.

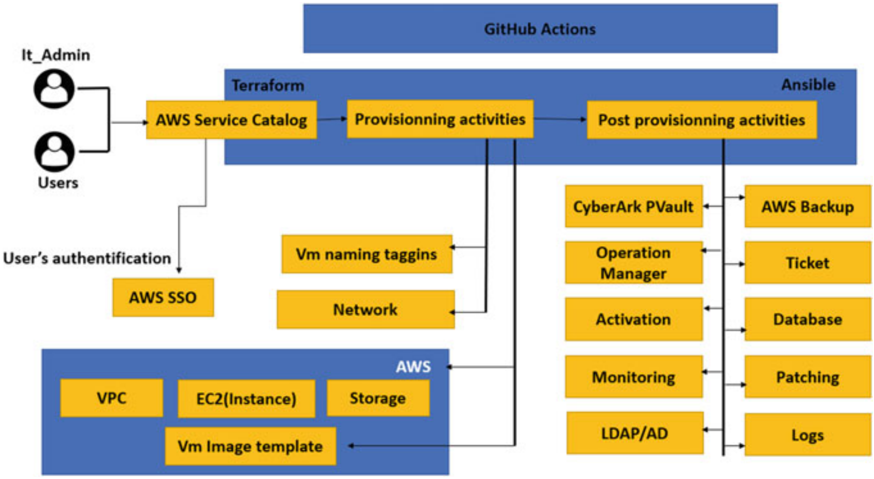


Fig. 5 Vm provisioning

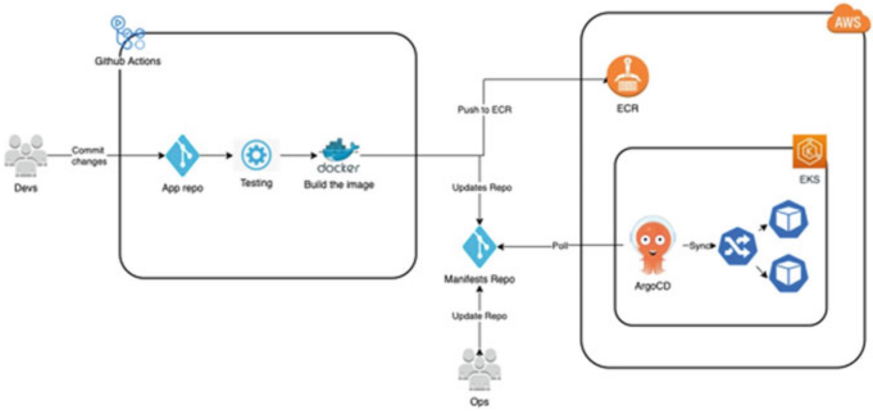


Fig. 6 Eco-system flows

All kubernetes manifest files representing objects will reside in the CLUSTER STATE repository in GITHUB. The cluster management operation of the pod's resources will be performed with the KUBECTL command [10] (Fig. 6).

1. The App developpers will commit their changes to the App repository
2. Using Github Action, the source code will be tested, the Docker image will be builed then pushed to EC
3. The pipeline will also update the manifests repository with the newest image version.
4. Other responsible teams can also push to the Manifest repository for Kubernetes cluster changes.

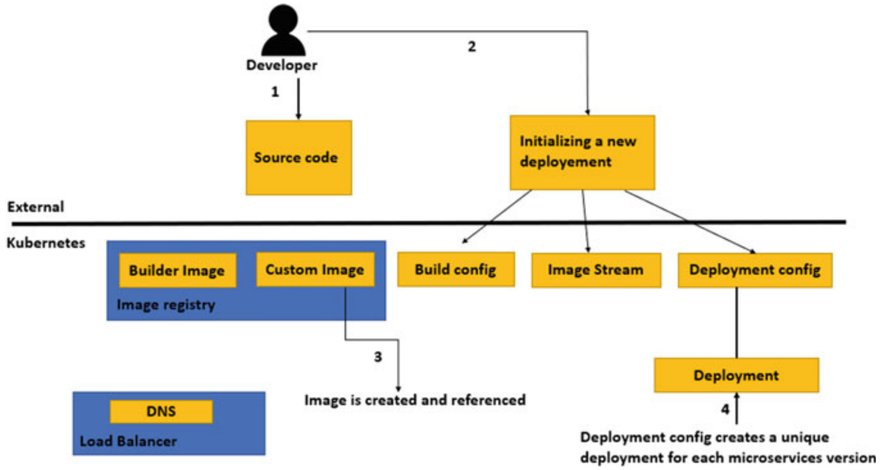


Fig. 7 Microservices flows

5. The Manifests repository will be the main source of synchronization for ArgoCD.
6. Argo CD will detect automatiquely new changes on the repository then synchronize Kubernetes cluster.

3.2 The Microservices Deployment Process

Figure 7 shows how these components are linked together. When a developer creates the source code and triggers a new application deployment, Kubernetes creates the components: deployment config, image stream, and build config [10].

1. Kubernetes developers creates a custom image using the source code and image template, this image is uploaded to the container image registry.
2. Kubernetes admins creates a build config to document the build of the application. This includes the image created builder and the source code location.
3. Kubernetes creates a deployment configuration, deploy, and update the applications. The Information contained in deployment configurations includes number of replicas, upgrade method, application-specific variables, and mounted volumes, more than that each unique application deployment is associated with the deployment configuration component.
4. The internal Kubernetes load balance is updated with an entry for the application's DNS record.
5. Kubernetes creates an image component. This image stream monitors the builder, if a change is detected, image streams will redeploy to reflect the changes.

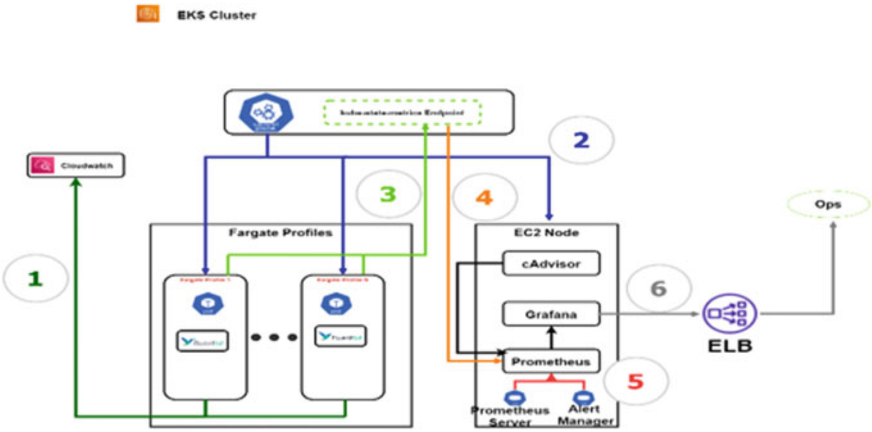


Fig. 8 Monitoring flows

4 Monitoring

The monitoring allows administrators to monitor all the resources on AWS account and the microservices inside the Kubernetes cluster (Fig. 8).

The schema below describes with more details this part:

1. Logging with aws cloudwatch service
2. AWS attach a fluenbit sidecar automatticaly to pod to retrieve logs and push them to cloudwatch
3. Configuration of the log router via the aws-logging configmap
4. Logging is automatically activated
5. For control plane, the logging is done by enabling it in the EKS creation phase
6. Control plane actively monitor and retrieves the state of kubernetes objects (Pods, deployments, Services ...) (2)
7. Kube-state-metrics service listens to the API server of the control plane and generates metrics about the state of the objects, these metrics are exposed by the API server in the kube-state-metrics endpoint (3)
8. cAdvisor generates resource usage and performance metrics of the running pods
9. Kube-state-metrics and cAdvisor exports the metrics to the Prometheus server (4)
10. Prometheus is stateful and needs a persistent volume of ebs type. So prometheus and grafana will be deployed on a NodeGroup (5)
11. Grafana expose dashboard for metrics gathered by prometheus through a load-balancer service wich should be accessible through https.

5 Conclusions

This article presents in depth a technical architecture for Hospital Information systems, based on IAAC tools and combined with cloud technology as IT service provider. A proof of concept, consisting the build of the global platform, based on these principals (IAAC, cloud) is presented and demonstrated. The combination of an infrastructure as a code, automation scripts and orchestration technologies for the deployment, the implementation and the configuration of the overall HIS platform without any human interaction.

As a future work we aim at using one dashboard to build and configure any solution based on the Cloud.

References

1. He C, Jin X, Zhao Z, Xiang T (2010) A cloud computing solution for hospital information system. In: 2010 IEEE International conference on intelligent computing and intelligent systems, vol 2, pp 517–520
2. Abdula M, Averdunk I, Barcia R, Brown K, Emuchay N (2018) The cloud adoption playbook: proven strategies for transforming your organization with the cloud. Wiley. [Online]. Available: https://books.google.fr/books?id=O1pQDwAAQBAJ
3. Boonchieng E, Duangchaemkarn K (2013) Application of cloud computing in the hospital drug information center in thailand. In: The 6th 2013 biomedical engineering international conference, pp 1–4
4. Yin J, Zhao D (2015) Data confidentiality challenges in big data applications. In: 2015 IEEE international conference on big data (big data), pp 2886–2888
5. Chen P, Freg C, Hou T, Teng W (2010) Implementing raid-3 on cloud storage for emr system. In: 2010 international computer symposium (ICS2010), pp 850–853
6. Imawan A, Kwon J (2015) A timeline visualization system for road traffic big data. In: 2015 IEEE international conference on big data (big data), pp 2928–2929
7. Santos J, Wauters T, Volckaert B, De Turck F (2019) Towards network-aware resource provisioning in kubernetes for fog computing applications. In: 2019 IEEE conference on network softwarization (NetSoft), pp 351–359
8. Nguyen TL (2018) A framework for five big vs of big data and organizational culture in firms. In: 2018 IEEE international conference on big data (big data), pp 5411–5413
9. Nishant Kumar Singh ST, Chaurasiya H, Nagdev H (2015) Automated provisioning of application in iaas cloud using ansible configuration management. In 2015 1st international conference on next generation computing technologies (NGCT), pp 81–85
10. Masek P, Stusek M, Krejci J, Zeman K, Pokorny J, Kudlacek M (2018) Unleashing full potential of ansible framework: university labs administration. In: 2018 22nd conference of open innovations association (FRUCT), pp 144–150

Computer Technologies in the Development of Quantitative Criteria for Calculating the Required Dose of Insulin in Patients with Type 2 Diabetes

Irina Kurnikova, Shirin Gulova, Natalia Danilina, Aigerim Ualihanova, Ikram Mokhammed, and Artem Yurovsky

Abstract ***Background***: Prescribing insulin to patients with type 2 diabetes may present as a problem even for endocrinologists, since there are no objective criteria for calculating the starting dose of insulin and the dose is often selected empirically based on the subjective assessment of a particular specialist; for example, by the selection method in a hospital with permanent observation: Set dose is prescribed and the level of sugar is monitored during the day. This is not possible in an outpatient setting. ***Methods***: Authors used a computer simulation method with 3D plotting to create a technology for calculating the required dose of insulin for patients with secondary insulin deficiency that developed with the background of type 2 diabetes. ***Results***: Clinical and laboratory examination of more than 200 patients with type 2 diabetes was conducted; the most significant correlations of the studied parameters (age, BMI) with the prescribed doses of insulin were identified; a formula for the calculation of the required insulin dose was derived with the construction of a 3D graph. ***Conclusions***: Obtained method allows to transfer patients with type 2 diabetes to insulin quickly and safely, which could be used both temporarily (for the period of surgery, illness, high levels of glycated hemoglobin) and as a part of combined or permanent insulin therapy.

I. Kurnikova (✉) · S. Gulova · N. Danilina · A. Ualihanova · I. Mokhammed
Department of Therapy and Endocrinology, RUDN University, Miklukho-Maklayast. 6, 117198 Moscow, Russia
e-mail: curnikova@yandex.ru

I. Kurnikova
Department of Aviation and Space Medicine, Federal State Budgetary Educational Institution of Further Professional Education, Russian Medical Academy of Continuous Professional Education, Barricadnaya St., H.2/1, B.1, 125993 Moscow, Russia

A. Yurovsky
Clinical Hospital of Civil Aviation, Ivankovo Highway, 7, 125367 Moscow, Russia

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_68

Keywords Type 2 diabetes · Insulin requirement · Secondary insulin resistance · Computer modeling

Background. Type 2 diabetes mellitus is a disease that in the first stages is accompanied not by insulin deficiency, but by excessive production of insulin. However, this does not exclude the formation of insulin demand at subsequent stages of the development of the disease, as a rule, after 5–6 years (constant overproduction leads to the depletion of insulin-producing β-cells) or insulin is prescribed for other reasons when the patient cannot intake medicine in the tablet form (presence of additional diseases), kidney disease, liver disease, accompanied by functional insufficiency, severe vascular and neuropathic complications of diabetes mellitus) or the level of glycated hemoglobin (an indicator of the quality of compensation for the last three months) is too high (above 9%).

Therefore, calculating the amount of insulin required for patients with type 2 diabetes is a difficult task, considering the aforementioned. For example, for a patient with type 1 diabetes, i.e., with absolute insulin deficiency, the required dose of insulin can be calculated considering the patient's weight (patient weight × 0.5 U/kg) or the amount of carbohydrates consumed (1–2 units of insulin per 12 g of carbohydrates or, as this amount of carbohydrates is also called, per 1 bread unit), but for patients with type 2 diabetes, these criteria are not suitable. There are other reasons for prescribing insulin to patients with type 2 diabetes—concomitant or, as they are also called, comorbid diseases; vascular complications of diabetes mellitus leading to functional failure of the kidneys and liver; neuropathic complications leading to the development of autonomic cardiac neuropathy and sudden death syndrome. Also, other factors of influence, such as an undetermined degree of preservation of residual secretion of insulin (depending on the individual characteristics of the organism), should be considered.

For the district physician, calculating the required dose of insulin is a difficult task, since they do not have experience in prescribing insulin therapy or in calculating the dose empirically. For the abovementioned reasons, it is now necessary to develop objective criteria (calculation formulas) for the dose of insulin for the treatment of patients with type 2 diabetes mellitus.

Study purpose: The aim of the study was to develop a measurement criterion using computer simulation to calculate the insulin intake in patients with type 2 diabetes mellitus at the beginning of insulin therapy with the background of a permanent insulin resistance.

Design: The study was conducted on the basis of the Endocrinology Department of City Clinical Hospital F.I. Inozemtsev (Moscow) in 2016–2018.

The study included 294 patients with type 2 diabetes, including 213 who received insulin therapy. All patients gave their informed consent to participate in the study.

Instruments and Data Collection Procedure

Clinical and laboratory examination was performed in accordance with medical and economic standards with an emphasis on medical and social characteristics, indicators of carbohydrate metabolism, the level of insulin resistance, and the required dose of insulin.

For statistical analysis, STATISTIC 10.0 computer program was used (Matematica®, Matlab®, HarvardGraphics®, StatSoft). The basic methods of statistical research were linear descriptive statistics (DescriptiveStatistics) with a calculation of the correlation of average standard deviations (corrs/means/SD).

Ethical Consideration: The research was approved by the Biomedical Ethics Committee of the Federal State Autonomous Educational Institution of Higher Education "Peoples' Friendship University of Russia", Protocol No. 9 dated March 17, 2016.

Results. The patients were divided into three groups (two observation groups, one comparison group): (1) observation group (104 people)—patients with type 2 diabetes receiving insulin therapy within the standard physiological requirement (up to 40 IU/day); (2) observation group (109 people)—patients with type 2 diabetes receiving insulin therapy more than 40 IU/day (secondary insulin resistance); and (3) comparison group (81 people)—patients with type 2 diabetes receiving oral hypoglycemic drugs (OHGD), see Table 1.

In patients of the observation and comparison groups, the quality of carbohydrate metabolism control (achievement of the target level of glycated hemoglobin—HbA1c) was assessed according to the Diabetes Control and Complications Trial (DCCT) criteria. In observation group 1, individual treatment goals were achieved in 15.1% of patients, in observation group 2—in 9.4%, in the comparison group—in 16.0% of patients.

In observation group 1, on average, the dose of insulin received was within the physiological requirement and amounted to 26.2 ± 4.1 U/day, and for patients in observation group 2, the daily dose of insulin was higher and was in the range of 73.7 ± 7.3 U/day. Some differences between the groups in terms of clinical features appeared at this stage. And if in patients in the first group the correlation between BMI and insulin dose is seen in more than 70% of patients, then in group 2 the correlation was even lower. In patients of observation group 2, there was also no clear relationship between BMI and the required dose of insulin, which confirmed the hypothesis of the influence of additional factors on the daily need for endogenous insulin.

It is well known that the production of endogenous insulin depends on body weight (BMI) and this trend persists in patients with type 2 diabetes mellitus. However, it seemed more significant in our study to explore the relationship between the level of exogenously administered insulin and BMI in our patients. And as our studies showed in observation group 1, where patients received an insulin dose of less than 40 U/day, there was no clear relationship between BMI and the dose of insulin received, which indicated that the residual secretion of insulin in the patient's body

Table 1 Medical and social characteristics of patients included in the study

Criteria	Patients receiving insulin therapy (observation group) ($n = 213$ people)		P_1	Patients receiving OHGD (comparison group) ($n = 81$ people)		P_2
	The target level of HbA1c was not achieved	Target HbA1c level achieved		The target level of HbA1c was not achieved	Target HbA1c level achieved	
Age	62.75 ± 4.6	63.41 ± 3.8	$p > 0.05$	59.12 ± 3.9	57.56 ± 5.2	$p > 0.05$
Gender (m/f)	0.77			0.88		$p > 0.05$
Duration of DM	9.99 ± 1.4	10.86 ± 2.0	$p > 0.05$	4.37 ± 0.9	5.69 ± 1.1	$p < 0.01$
BMI	32.31 ± 3.8	29.65 ± 2.4	$p > 0.05$	32.92 ± 4.4	30.24 ± 2.9	$p > 0.05$
Duration of insulin intake	3.36 ± 1.0	5.05 ± 0.8	$p < 0.05$	–	–	
The number of patients who studied at "School of Patient with DM"	46.9%			38.6%		$p < 0.05$

Remark HBA1c glycated hemoglobin; *BMI* body mass index; *OHGD* oral hypoglycemic drugs; P_1 significance of differences within the group; P_2 significance of differences between groups, *DM* diabetes mellitus

was sufficient. We verified the obtained data using the computer simulation method—surface plotting. The correlation between BMI and the required dose of insulin did not have a pronounced dependence, but the duration of the disease was important (Fig. 1).

We obtained confirmation by constructing similar graphs for subgroups (observation group 1 and observation group 2). In patients of observation group 1 (with preserved production of their own insulin), there was no clear dependence of the insulin dose received on the duration of the disease and on the BMI, but the strong relationship with the age of patients was confirmed. Age-related insulin resistance is a well-known phenomenon and was confirmed by qualitative studies by Peters et al., as early as 1989. But in patients who needed a dose higher than the physiological need to compensate for carbohydrate metabolism (observation group 2), the value of this dose depended on both the duration of the disease and BMI. The obvious reason for this dependence was a decrease in the production of one's own insulin and dependence on insulin income from outside. In this case, standard mechanisms began to work to ensure the need for insulin per kilogram of body weight and the impact of a physiological decrease in insulin sensitivity in older age groups. Predicting the formation of secondary insulin resistance is at the same time an indicator of the assessment of a decrease in the production of one's own insulin. And the study of

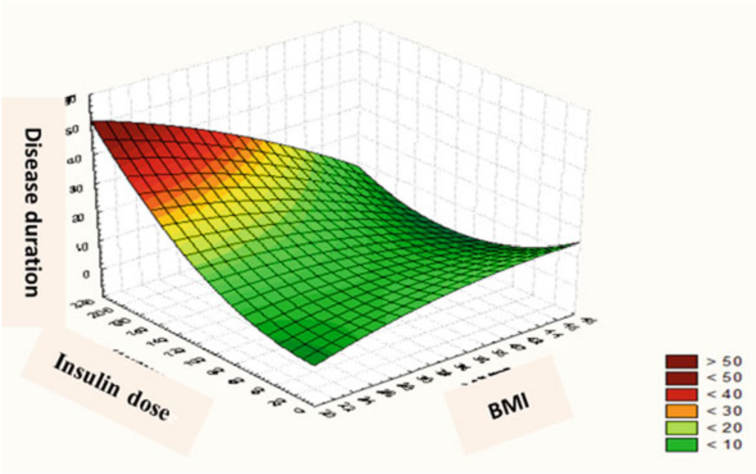


Fig. 1 Three-dimensional surface plot of insulin dose versus BMI and disease duration

the factors influencing this process allows medical professionals to make a timely forecast.

Significant criteria for determining the required dose of insulin were: the duration of the disease, which reflected the rate of decline in own insulin production in years, and the body mass index, which reflects the sensitivity of insulin receptors.

This calculation formula was obtained on an array of 213 patients with type 2 diabetes already receiving insulin therapy. From that array, 42 patients were identified as satisfactory compensation for diabetes mellitus receiving insulin therapy. For each of these 42 patients, three criteria were determined: duration of disease, in years (DD); BMI (kg/m^2); dose of currently received insulin (unit). Variation series were constructed by age, duration of diabetes mellitus, body mass index, and daily dose of insulin in the examined patients with a satisfactory quality of diabetes compensation. Based on the obtained variation series, a three-dimensional graph was constructed by the method of multiple linear regression (Fig. 1), where the duration of the disease in years was plotted along the X axis, BMI was plotted along the Y axis, and the dose of insulin received was plotted along the Z axis.

As a result, a Formula (1) for determining the required dose of insulin (RDI) per day was obtained using the following derived formula:

$$\mathrm{RDI} = -56.7 + 3.2 \cdot \mathrm{BMI} - 0.2 \cdot t, \tag{1}$$

where RDI—required dose of insulin per day; BMI—body mass index; t—duration of diabetes in years; − 56.7 and 3.2 and 0.2—numerical values of coefficients. The invention allows to quickly determine the required dose of insulin in type 2 diabetes mellitus, which is relevant in solving this problem.

56.7—numerical value subtracted from the difference of the obtained products of factors (3.2 × BMI) and (0.2 × DD), calculated using the computer simulation method when constructing 3D graphs in the analysis of 213 clinical cases.

3.2 and 0.2—the calculated coefficient obtained by constructing the multiple regression equation between BMI, duration of the disease, and the dose of insulin received.

The required or estimated dose of insulin was assessed by the quality of diabetes compensation in terms of glycemia and glycated hemoglobin, and the calculated value of the insulin dose is sufficient for adequate control of blood glucose levels (glycemia indicators are within the target values for a particular patient).

A patent for the invention RU 2 684 393 C1 "Method for determining and calculating the required insulin dose in type 2 diabetes mellitus and established insulin dependence" was obtained on August 31, 2018 [Kurnikova Irina Alekseevna (RU), Aigerim Ualihanova (KZ)] [1].

Conclusion: Constant hyperstimulation depletes the resources of pancreatic beta cells, and over time, the relative insufficiency of insulin (the lack of its entry into the cell) becomes absolute (a decrease in the production of insulin). And at later stages, after 5–7 years from the onset of the disease, the result of prolonged hyperstimulation of pancreatic β-cells that produce insulin is their depletion, and relative insulin deficiency begins to turn into absolute insulin deficiency.

At this stage, a patient with type 2 diabetes already needs insulin therapy (insulin injections) or combination therapy (insulin injections + tablets). Methods for calculating the required dose of insulin in this case have not been developed.

In cases where the patient develops a need for insulin—the main problem in selecting the dose of insulin when transferring from glucose-lowering drugs to insulin therapy is determining the adequate dose of the drug. Very few methods are generally used to promptly switch to insulin in the preoperative (elective or emergency) phase of patients with type 2 diabetes. For example, a method based on calculating the dose of insulin according to the level of hyperglycemia [2].

The dose of insulin is calculated according to the formula: insulin units/h = blood glucose concentration (mg%)/150. Or another method for correcting hyperglycemia in patients with diabetes mellitus, based on the selection of insulin doses according to an increase in blood glucose concentration in response to a standard pain stimulus (subconjunctival injection) [3].

These methods are provided only for use in emergency situations and do not allow calculating the insulin requirement in conditions of normoglycemia or hypoglycemia, and with painful stimuli, there is even a risk of developing vasospasm in response to the release of "pain hormones", which can lead to coronary or cerebral circulation disorders.

The use of computer technology makes it possible to solve this problem promptly and does not require the use of invasive methods. The proposed formula for calculating the required insulin dose in type 2 diabetes patients with reduced insulin secretion provides the ability to calculate a physiologically reasonable dose of insulin under conditions of normoglycemia, hyperglycemia, and hypoglycemia.

References

1. Patent for invention RU 2 684 393 C1 "Method for determining and calculating the required dose of insulin in type 2 diabetes mellitus and established insulin dependence" 31/08/2018. Bulletin of the Federal State Institution "Federal Institute of Industrial Property and the Federal Service for Intellectual Property, Patents and Trademarks", "Inventions, Utility Models", No. 10, 2019 https://edrid.ru/rid/219.017.0b28.html
2. Edward G (1996) Morgan Jr. Maged S. Mikhail clinical anesthesiology, 2nd edn. Stamford, p 882
3. Maksimov VYu (1990) Prediction and prevention of complications during cataract extraction in patients with diabetes mellitus: Abstract of diss cand med Sciences. Kuibyshev, p 23

Security of Input for Authentication in Extended Reality Environments

Tiago Martins Andrade, Jonathan Francis Roscoe, and Max Smith-Creasey

Abstract In this concept paper, we evaluate the security impact of accelerometer data for authentication in extended reality (XR) environments. Currently, there is a lack of authentication mechanisms in VR/XR environments. Most authentication is carried out through PINs and passwords which detracts from the immersive experience and inconveniences the user. Motion-based gesture techniques have recently shown potential in authentication users in VR environments. However, the state-of-the-art works have not considered the issue of VR being a visible activity which would yield gestures used to authenticate vulnerable to mimicry. We demonstrate how subtle changes to a user interface (UI) can increase the complexity and cost of eavesdropping on users in VR environments and propose directions for future research. We call on the industry to acknowledge and design around the unique security challenges of authentication in VR.

Keywords Virtual reality · Authentication · Application security · Biometrics (access control) · Keystroke dynamics · Computer security

1 Introduction

Virtual reality (VR) and extended reality (XR) systems are very quickly becoming a mainstream and powerful technology, with the market expected to reach a total value of almost 300 billion US dollars in 2024 [1]. Whilst the use cases for VR/XR technologies thus far have largely been recreational, there are developments in applications for VR headsets that require a level of security (such as virtual security operations centres (VSOCs) [5]). Such environments provide access to privileged information and therefore need a stringent level of authentication to keep non-authorised users out of the system. Insufficient authentication and authorisation mechanisms within a secure VR environment could have significant implications for operational security.

T. M. Andrade · J. F. Roscoe (✉) · M. Smith-Creasey
BT Applied Research, Adastral Park, UK
e-mail: jonathan.roscoe@bt.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_69

In order to protect against use from non-authorised users, many systems use authentication techniques such as passwords in which a user must use Bluetooth-connected controllers to input their password into a virtual keyboard or follow specific steps to unlock certain content. However, if the system has been compromised, and the attacker is able to store all user movements, it is possible to trace all user steps one by one in a simulated environment within VR/XR. For example, if the user is writing down a password using a virtual keyboard, by mimicking all user movements, and since the virtual keyboard is static, it is possible to extract the exact password and gain unauthorised access.

Users of XR are particularly vulnerable to eavesdropping on interactions due to their lack of awareness of their surroundings. There is a potential approach to attacks from visual observation as well as captured accelerometer data that could lead to password mimicry.

This concept paper proposes and conducts an investigation into the level of gesture robustness and the possibility of obfuscating that data from a mimicry attack with simple UI changes. We compare different approaches for the virtual keyboards (original layout, control layout, adjusted layout, and randomised layout). We hypothesise that the randomisation of the entire keyboard layout will degrade the usefulness of the accelerometer data extracted from user movements.

1.1 Related Work

Most commonly, authentication is split into (i) something the user *has*, (ii) something the user *knows*, and (iii) something the user *is*. The password remains the most common form of authentication today despite it often leaving users fighting against security for usability. Within the VR space, some authentication has used biometrics but the most prevalent form of authentication in VR systems today still only rely on the password [6]. However, some studies have found that the combination of knowledge *and* biometric information can yield better security [7]. The interaction with VR (such as the input of a password (or any text)) carries additional challenges because attackers that are able to collect the accelerometer data during input might be able to make inferences about the interaction [2].

The previous work has explored unconventional approaches to acquiring data such data surreptitiously, including human activity and video [8]. Consequently, collection of accelerometer data from a smartwatch or fitness wearable may be a viable attack mechanism for XR users [2]. Despite this research, there is a lack of investigation into the vulnerabilities such side-channels pose to user typing input (e.g. a password) or possible solutions towards the mitigation of these side-channels.

2 Proposed Approach

2.1 General Idea

We can see from existing research that if a VR system or a wearable is accessed by an attacker that extracts the accelerometer data that could allow the attacker to re-create activities [2]. Therefore, we hypothesise that slightly randomised changes (with differing levels of granularity) to the UI or 3D objects could be sufficient to obfuscate user actions on that data and at the same time not increase complexity or extra steps to the end user. Therefore, an experimental study to extract that information was conducted.

2.2 Implementation

To assess and collect user movements when interacting with a virtual keyboard, a VR application was created using Unreal Engine 4 [4]. The head-mounted display (HMD) used in the experimental study was Oculus Quest connected to a computer by Oculus Link and the use of the Oculus Motion Controllers.

In the design process, to make virtual reality to be effective, it is important to fulfil the 3 illusions basic principles [9], to assure that the user is immersed in the experience and all user perceptions match reality or at least the user expectations of an certain action/reaction. The 3 following illusions which need to be in place are:

– **Place Illusion**: The feel of being in a virtual place, even though you know you are not there;
– **Plausibility Illusion**: Illusion of the perceived events to feel real for the user;
– **Body Ownership**: Your virtual body is connected to your body.

A simple virtual environment was created with a floor, a sky dome, and default light/shadows. When the user starts the VR application, they are placed in the middle of that environment and presented 8 cubes with letters as shown in Figs. 1, 2, and 3. This acts as a simplified keyboard where the user is tasked with writing a single word multiple times.

The keyboard keys can change shape, size, and location during specific events to allow capture of the movements data to be analysed. There are 4 possible changes for the keyboard:

1. **Original layout**: This layout is static and predetermined by us (as shown in Fig. 1), and it will be always the same during sessions and for the entire experimental period.
2. **Control layout**: This layout is static and a full copy of the original layout (as shown in Fig. 1). This will allow us to compare the same layout and verify if the movements will match (be the same) when using a static keyboard layout.

Fig. 1 Fixed key layout

Fig. 2 Adjusted key layout

3. **Adjusted layout**: This layout is static but the keys will randomly change places at the beginning of the session for each user(as shown in Fig. 2).
4. **Randomised layout**: This layout is completely randomised. The key location, size, and shape will randomly change at the beginning of each session and will be always different for each user (as shown in Fig. 3).

For each of these keyboard layouts, the user is tasked with entering the word 'PILOT' ten times. To do this, they point with their dominant hand at the appropriate

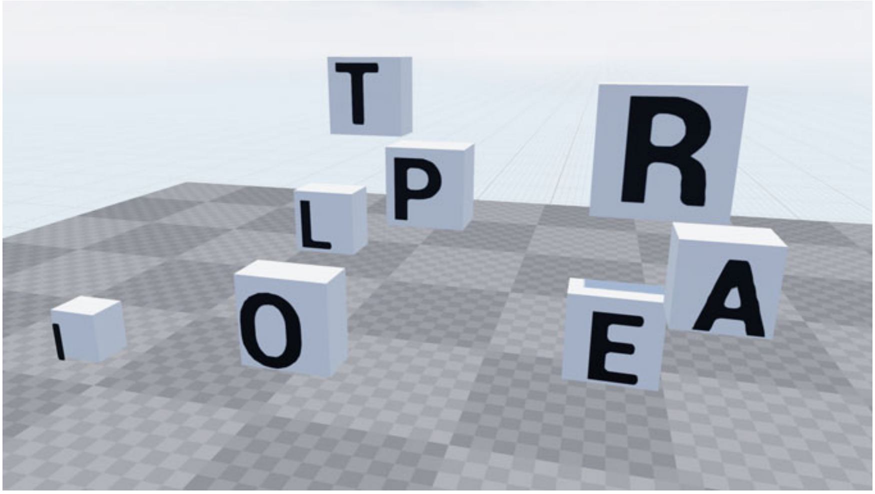


Fig. 3 Random key layout

cube and click the trigger button. Whilst the user is typing the word, the angular acceleration of the active controller is constantly logged.

3 Experimental Results and Discussion

To visualise result, we calculate the magnitude of angular acceleration measured during the authentication process. Angular acceleration is a three-dimensional vector of angular acceleration in rad/s^2. We calculate magnitude (m) with:

$$m = \sqrt{x^2 + y^2 + z^2}. \tag{1}$$

Figure 4 visualises the magnitude of angular acceleration for all users, from which we can see the original and control layouts have significantly less total motion. In Fig. 5, the mean magnitude for all samples is split by the input method and clearly shows that the distribution is much greater for adjusted and randomised input methods.

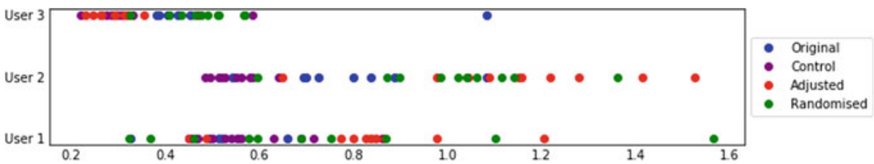


Fig. 4 Mean acceleration magnitude for different layouts, grouped by user

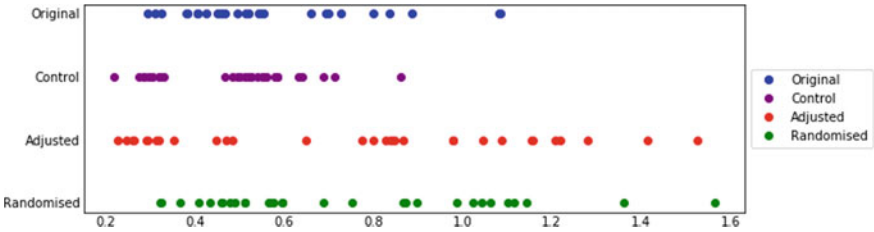


Fig. 5 Mean acceleration magnitude for all users, grouped by keyboard layout

It is notable that although there is a clear difference between the original and adjusted layouts, there is only a slight change in distribution for adjusted and randomised. This suggests that despite significant changes, the users are not making significant adjustments with their hand position, increasing the size of the field (i.e. utilising the 360° space may enhance this).

4 Conclusions and Future Work

Our initial results demonstrate that a UI with fixed layout results in a predictable range of motion, we posit that this indicates a potential for eavesdropping on an XR environment. We suggest that UI design should be carefully considered and may benefit from an element of randomness as standard practice.

The randomised layouts prevent a user from carrying out identical gestures and increasing the amount of motion when entering identical data, we propose a further range of adjustments to a virtual environment to introduce noise into eavesdropping attempts [3]. These can be tailored to have subtle impact on detectable user motion.

For the purposes of access, a further counter-measure could be the deployment of continuous authentication in XR environments that can constantly validate input from a user based on biometric characteristics. Another potential approach to mitigating some forms of eavesdropping is to increase user awareness of the external environment and potential malicious observers.

In the future work, we wish to explore more fully the ability of a malicious observer to predict input with varying degrees of knowledge and conduct mimicry attacks, to understand the level of information extraction that may be achieved. With regards to the design of secure interfaces, we would like to experiment the spread of user input over a full 360° sphere.

References

1. Alsop T (2021) Augmented reality (AR), virtual reality (VR), and mixed reality (MR) market size worldwide in 2021 and 2028. https://www.statista.com/statistics/591181/global-augmented-virtual-reality-market-size/
2. Andrade TM, Smith-Creasey M, Roscoe JF (2020) Discerning user activity in extended reality through side-channel accelerometer observations. In: 2020 IEEE international conference on intelligence and security informatics (ISI), pp 1–3. https://doi.org/10.1109/ISI49825.2020.9280516
3. Andrade TM, Smith-Creasey M, Roscoe JF (2022) Security method for extended reality applications. Patent GB22003313 Jan 2022
4. Epic Games: Unreal engine (2019). https://www.unrealengine.com
5. Hercock R (2021) Why AI is here to stay in cyber defence. https://www.globalservices.bt.com/en/insights/blogs/why-ai-is-here-to-stay-in-cyber-defence
6. Jones JM, Duezguen R, Mayer P, Volkamer M, Das S (2021) A literature review on virtual reality authentication. In: International symposium on human aspects of information security and assurance. Springer, pp 189–198
7. Mathis F, Fawaz HI, Khamis M (2020) Knowledge-driven biometric authentication in virtual reality. In: Extended abstracts of the 2020 CHI conference on human factors in computing systems, CHI EA '20. Association for Computing Machinery, New York, NY, USA, pp 1–10. https://doi.org/10.1145/3334480.3382799
8. Roscoe JF, Smith-Creasey M (2020) Unconventional mechanisms for biometric data acquisition via side-channels. In: 13th International conference on security of information and networks, pp 1–4
9. Slater M (2017) Implicit learning through embodiment in immersive virtual reality, chapter 1. Springer Singapore, Singapore, pp 19–33. https://doi.org/10.1007/978-981-10-5490-7_2

Showing the Use of Test-Driven Development in Big Data Engineering on the Example of a Stock Market Prediction Application

Daniel Staegemann, Ajay Kumar Chadayan, Praveen Mathew, Sujith Nyarakkad Sudhakaran, Savio Jojo Thalakkotoor, and Klaus Turowski

Abstract The concept of big data has huge implications for today's society and promises immense benefits if used correctly, but the corresponding applications are very error-prone. Therefore, testing must be as comprehensive and rigorous as possible. One of the solutions proposed in the literature is the test-driven development (TDD) approach. TDD is a software development approach that has a long history, but has not been widely applied in the big data domain. Nevertheless, a microservices-based TDD approach has been proposed in the literature, and the feasibility of its application in actual projects is studied here. To this end, a stock market forecasting application is implemented as an exemplary use case. It comprises seven microservices, an additional database, and the connection to an external service. However, the focus is explicitly on the TDD of the services and their interaction. The actual quality of the forecasts is only a secondary aspect with little relevance to the presented research.

D. Staegemann · A. K. Chadayan · P. Mathew (✉) · S. N. Sudhakaran · S. J. Thalakkotoor · K. Turowski
Otto-von-Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany
e-mail: praveen.mathew@ovgu.de

D. Staegemann
e-mail: daniel.staegemann@ovgu.de

A. K. Chadayan
e-mail: ajay.chadayan@ovgu.de

S. N. Sudhakaran
e-mail: sujith.nyarakkad@ovgu.de

S. J. Thalakkotoor
e-mail: savio.thalakkotoor@ovgu.de

K. Turowski
e-mail: klaus.turowski@ovgu.de

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_70

Keywords Big data · Test-driven development · Testing · Microservice · Software engineering · Quality assurance

1 Introduction

With the increasing significance of data and their processing and, thereby, also the concept of big data (BD), the assurance of the corresponding applications' quality also gains importance. One rather recent proposition for this challenge was the application of test-driven development (TDD) to the development of BD applications [1]. For this purpose, it was suggested to utilize microservices [1]. Their main goal is to enable loosely coupled, self-contained modules or services that are created to solve a specific task, have their own resources, and can be deployed separately. Various asynchronous communication patterns can be used between services and messaging. For instance, the event-driven approach and RESTful connections are some widely used patterns in software engineering. Since microservices are independent components, it is also feasible to use different programming languages for their implementation [2]. Moreover, they have already proven to be a valuable tool in a BD context [3].

While the application of TDD in BD has already been demonstrated [4, 5], the related literature is still relatively sparse. This publication therefore aims to extend it by presenting an additional use case, more specifically a stock market prediction application. However, the focus is explicitly on the TDD of the services and their interaction. The actual quality of the forecasts is only a secondary aspect with little relevance to the presented research.

2 Test-Driven Development

Based on the corresponding scientific literature [6], TDD can be characterized as a way to improve an implementation's quality for the cost of increasing the development time and associated effort. However, depending on the use case, this trade-off might be more than worth it, making it a valuable part of a developer's tool kit.

This approach of developing software aims at improving its quality by mainly addressing two factors. The first one is the expected increase in test coverage. Consequently, with more of the code being tested in a meaningful way, it is expected to also identify more issues and bugs. These can subsequently be fixed, which improve the developed software's quality. Additionally, the software's design process itself is also influenced. The inherent focus on small, incremental steps usually leads to a better planned and better manageable structure, which in turn makes it easier for the developers to avoid bugs and incompatibilities [7, 8]. While TDD is mostly used in software development, the application in but process modeling [9], the special case of implementing BD applications [1], and developing ontologies [10, 11] have also been discussed in the literature.

When following the “traditional” software development paradigm, a function or a change that is to be realized is first implemented and afterward tested. In contrast, TDD reverses the order of implementation and testing. That is, after the desired change is designed, it is segmented into its smallest meaningful parts [12]. Then, one or more tests are written for these. These aim to ensure that the desired functionality is provided. Afterward, the tests are run. However, they are expected to fail because the actual functionality is not yet implemented [13]. Only after this step, the code that actually provides the new functionality is written. At this stage, factors beyond the pure functionality are ignored. This includes, for instance, the elegance of the code. Instead, the simplest solution is pursued. When the functionality is implemented, the code must pass the tests that were written beforehand [7]. If no issues are found, the code is revised with regard to factors such as readability or compliance with standards and best practices [13]. In this process, the tests are constantly utilized to validate the code.

Yet, applying TDD affects more than just the test coverage. It also influences the software design because instead of large tasks, small work packages are used. Furthermore, this emphasis on incremental changes [14], interviewing the testing and implementation to short test cycles, provides the developers with immediate feedback [15]. Even though most tests are specifically created for these small units, other types of tests such as acceptance, integration, or system tests can also play a role in TDD [16]. To fully harness the potential of TDD without tying up developers’ attention by forcing them to execute tests manually, TDD is often used in conjunction with test automation as part of continuous integration (CI) and continuous development (CD) efforts [17, 18]. To make sure that the latest amendments to the code do not cause issues for already existing parts of the implementation, upon the versioning system registering a new code commit, a CI server automatically starts and re-executes all applicable tests.

3 The Implementation

In the following, the developed application is described, comprising the general architecture as well as the individual services. Furthermore, it is outlined how the TDD was implemented.

3.1 The Application Architecture

The stock prediction application consists of seven microservices. There is one frontend service and there are six backend services. An overview, also including the utilized database and an external service as information source, is given in Fig. 1.

The services communicate directly with each other using HTTP requests/response protocol. In order to use the scheduler, we use java ScheduledExecutorService, which

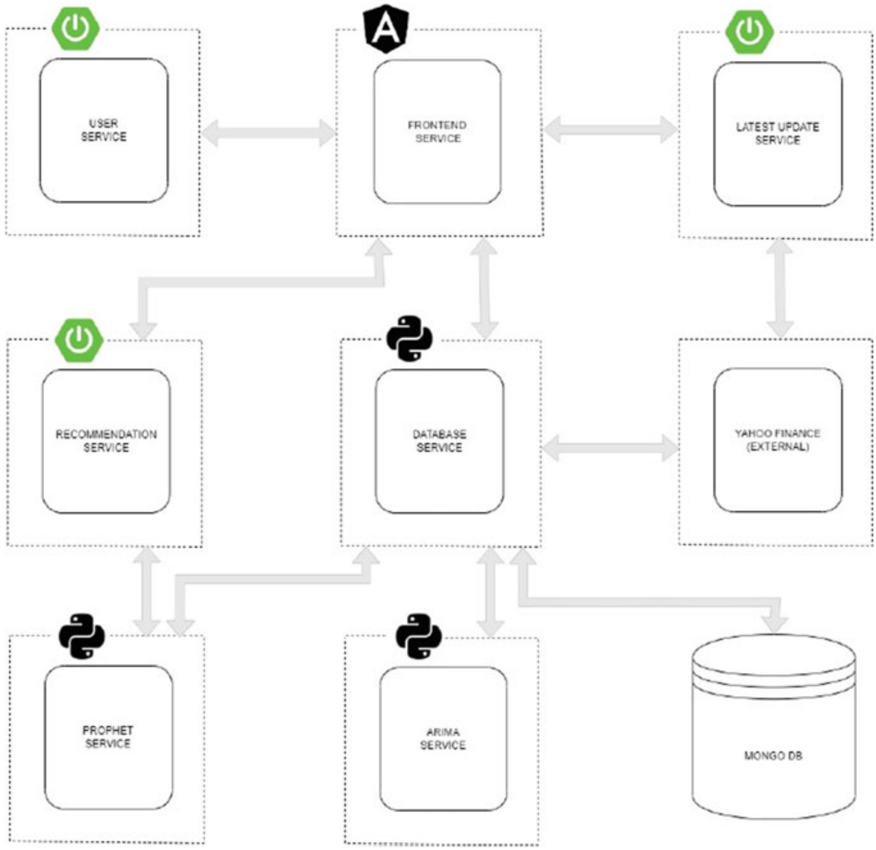


Fig. 1 Application architecture

is responsible for scheduling events to refresh data for reflecting the latest stock information. All the individual service communications are outlined in Table 1.

In the following, the services are briefly described, to provide the reader with a general understanding of their functionality and interplay and to also give some context to the later on following explanations concerning their testing.

Frontend Service The frontend service was initially built using the Flask framework. It is a microweb framework written in Python. The main advantage of using Flask was the built-in development server and the fast debugger provided. Even though the lightweight flask had many advantages in the TDD approach, it had limitations in testing. To apply the TDD in the frontend, Flask was replaced with Angular framework. Angular is a TypeScript-based web application framework. Angular has a component structure and uses a combination of Typescript, Html, CSS, and Typescript spec test files. As each component has its typescript and corresponding test file, following TDD was much easier and more efficient with Angular.

Table 1 Communication between the services

Communications	Description
Frontend service ↕ User service	For all the user login/signup data, the frontend service communicates with the user service which holds all the data
Frontend service ↕ Latest update service	The frontend service communicates with the latest update service to determine the current top three performing companies. The latest update service fetches the latest stock data from Yahoo Finance
Frontend service ↕ Prophet service ↕ Database service	The frontend communicates with the prophet service, which in turn fetches the data from the database and produces the prediction results. These results are further returned to the front end and displayed to the user
Frontend service ↕ ARIMA service ↕ Database service	The frontend communicates with the prophet service, which in turn fetches the data from the database and produces the prediction results. These results are further returned to the frontend and displayed to the user
Frontend service ↕ Recommendation service ↕ Database service ↕ Prophet service	To get the recommendation for the next week and populate the graphs, the frontend service communicates with the recommendation service which internally communicates with the database service to run prediction service and get the best-performing company for the upcoming week

The application consists of a service layer named API-service, which handles the HTTP requests of the application. The service layer communicates with the components like signup, login, and dashboard, which is the layer responsible for the business logic.

There are mainly three components named dashboard, login, and sign up. The login and the signup components deal with the login/signup of the user and connect to the user service in the backend. These services solely consist of business logic related to the users. The dashboard component communicates with three services, namely backend service, recommendation service, and update service. The dashboard component is segregated into three parts to demonstrate the data received from these three services. All components of the application are built following the single responsibility principle from the SOLID principles to keep the code clean and understandable.

Stock Prediction Service (Fbprophet) The first service for predicting future values of stocks was created based on a software named Prophet. It is an open-source software from Facebook's Data Science Team that has a procedure designed to forecast time-series data based on an additive model where nonlinear trends are fit with yearly, weekly, and daily seasonalities including the holiday effects. The library is known for generating high-quality time-series forecasts where the procedure works best with time-series data with strong seasonal effects with several seasons of historic data [19, 20]. Prophet was chosen because it is fast, has the ability to provide highly accurate forecasting, has minimal requirements for data preprocessing, and is robust toward missing data and outliers. In our implementation, the data are loaded into a Pandas data frame, and the model is trained and is then used to create future data frames and forecast forward.

Stock Prediction Service (ARIMA)

ARIMA is a statistical model for analyzing and forecasting time-series data. ARIMA models are considered the most robust and efficient in forecasting financial time series, especially for short-term prediction [21]. The parameters of the ARIMA model were calculated using the "auto_arima" function. ARIMA was chosen as the second option for the prediction services because in the past it constantly outperformed various complex models in short-term prediction [21]. In our implementation, the data are loaded into a Pandas data frame, then the "auto_arima" function is used to calculate the parameters for the ARIMA function, and subsequently, the data are passed through the model to forecast the stock performance for future dates.

Yahoo Finance and Database Service YFinance is an open-source Python-based library used for downloading stock prices of different companies [22–24]. In our project, we have used this library to download stock data of companies such as Apple, Amazon, Google, Facebook and perform computations on the extracted data.

In order to connect Python to the utilized MongoDB database, we used pymongo, which is an open-source library for Python Mongo database connectivity. We hosted our database server using MongoDB compass which is a service provided by MongoDB to host the database on a remote server. This free hosting service helps us in creating better availability and scalability as it follows the master–slave architecture with master and slave nodes forming clusters. In our implementation, the data of the relevant companies are downloaded using the yfinance get_stock_info() service. The data are then loaded into a pandas data structure and subsequently persisted into the database through the "insert_data()" method of the database service.

Latest Update Service The latest update service is a microservice built using the Spring Boot framework in Java. The library dependencies are handled using maven. This microservice is used to determine the top three performing companies for the current date. The application consists of a controller that handles all the HTTP requests to the application and a service layer that carries out all the business logic. Below is a diagram that illustrates the architecture of the latest updated service. Here, the application of TDD helps to resolve linting errors which are caught using the Checkstyle Plugin. A custom Checkstyle rule file was added to monitor the quality of the code. Checkstyle was configured in maven and will automatically run while a merge request/push to the main is performed with help of GitHub Actions. While checkstyle checks the presentation of the code, SonarQube was used to monitor any code vulnerabilities or bugs undetected by unit tests.

User Service The user service is a cache-based microservice built using the Spring Boot framework in Java. The library dependencies are handled using maven. This service handles all the requests with respect to the user and the business logic related to the user like the addition of new users and login authentication details. The application is split into various packages to maintain abstraction.

The user controller class handles all the HTTP requests to the application, and it internally communicates with the repository through the factory and the model classes. The model classes hold the object and are mapped as an entity to the database. The application hosts a database to hold all the details. The properties for the database are configured using an application properties file. To enhance the performance, spring boot caching was implemented in the application at the controller level. The list of users is cached on the first request and is afterward returned from the cache. When a new user is added, the cache is evicted. By enabling caching, it was possible to reduce the time required to answer an exemplary request from 686 to 7 ms. As for the "Latest update service", checkstyle and SonarQube were used to control the codes quality.

Recommendation Service The recommendation service is a microservice built using the Spring Boot framework in Java. The library dependencies are handled using maven. This microservice is used to deliver the stock recommendation. It internally communicates with the backend service and runs a prediction algorithm that determines the expected best-performing company and returns the values. The application consists of a controller that handles all the HTTP requests to the application and a service layer that carries out all the business logic. As for the "Latest update service" and the "User service", checkstyle and SonarQube were used to control the codes quality.

3.2 The Testing

In general, Pytest and UnitTest were used to write our test scripts for those services written in Python. Furthermore, Mockito and Junit were used to test services written with Java. While not every single test can be outlined, in the following subsections, some exemplary tests are highlighted.

To automate the testing when changes are implemented, GitHub actions were used to create a CI/CD pipeline for the stock prediction application. A YAML file was added to the repository which triggered the action every time code which was pushed to the main branch or a merge request was created to the master. Currently, only the java projects are added to the CI/CD pipeline. The build starts with setting up the JDK for the projects and then runs the unit tests as well as the checkstyle checks for the projects. If all the tests pass, a Docker image is created for the service, and the image is pushed to the repository in Docker Hub.

Stock Prediction (fbprophet) The requirements for the prediction service are thoroughly based on the actual requirements for the prophet library to work. The algorithm within the library requires data about the stock in a specific format, the number of days for the prediction from the user, and furthermore, the output of the prediction should be in the desired format so that it can be fed to the services at the front end. Therefore, for this prediction service, three functions were needed, namely a preprocessing function, the prediction function, and a postprocessing function.

The data have to be preprocessed according to the requirement of the prophet, which is to have an input train containing only two columns—the date and the closing values. The column names should also be specific. Therefore, there are two checks in place. Firstly, the preprocessed data have only the date and the corresponding closing values. Moreover, the names of the columns are passed and have to be “ds” for the date and “y” for the closing value.

The forecast function can be written based on tests of the type of its output. According to the prophet documentation, the forecast output will contain nineteen columns. As the input is passed on in a Pandas frame, the output expectation would also be the same. Therefore, it is checked if the output is in the pandas frame and if the result has a total of 19 columns.

Stock Prediction (ARIMA) The algorithm within the library requires the following three properties. Firstly, data about the stock need to be in a specific format. The input data are sourced from the database where data from the Yahoo Finance webpage are stored and should contain seven columns and should be in data frame format. Further, the ARIMA model should only be used when the accuracy is more than 75%. Lastly, the results are checked and it is confirmed that we at least have two columns, one containing the future dates and the other having values, for those days.

Regarding the first one, the data have to be validated so that it is available in a specific format that is easier for the algorithm to work on. Hence, it is checked if the input is in the Pandas frame and if it has a total of seven columns. Since multiple algorithms are being used, it was decided that we would consider the ARIMA algorithm only if it gives us an accuracy above the desired threshold, which is 75% in this case. Consequently, this is evaluated. Finally, the data that have been received from the algorithm have to be checked for the format so that it can be used for plotting. Therefore, it is checked if the output has two columns, the future dates, and their corresponding values.

Frontend Service The requirement for the frontend service was to host a UI for the user to select the company and to show the prediction data. To follow the TDD approach, the frontend service was migrated from the initially chosen flask to angular. With an angular framework, each page could be created as a component and each could be tested individually. Tests were written to check if buttons on the frontend were clicked and intended functions were called on each activity. Events like onclick, ngOnInit were tested to check if the variables had a value or were undefined.

User Service The requirement for the user service was to process the user login/ signup details. To achieve it, a database was created to hold the data. The user service communicated with the frontend service and tests were initially written to test the connections. As this was a cache-based microservice, a test for the cache was also included. Further, because the service dealt with data from the database, tests for data retrieval from the database repository as well as data insertion were added.

Latest Update Service The requirement of the service was to provide the frontend with the top three performing companies for the present week. To determine this, the service fetched the present data from Yahoo finance and returned the top three companies. Tests were written to check if the controller returned values to the frontend and if the services returned the stock data after processing. The tests were able to cover 100% of the lines in the controller and about 81% in the service.

4 Discussion

The presented research aimed to explore and demonstrate how to effectively utilize the TDD approach to implement a big data application. The tests were created for frontend and backend implementations. Challenges were mainly faced during the initial setup phase. We also developed a CI/CD pipeline, which enabled autobuild and deployment if the test cases are successful. TDD for a normal microservice application is very efficient when it comes to reproducing similar lines of code. However, it gets more complicated when using custom machine learning algorithms, because the output is hard to determine beforehand and hence it is also challenging to formulate the requirements. Yet, we used algorithms that were already designed and using them was considerably easier.

Further, for our implementation, we relied on several best practices. Proper naming conventions were followed, so all the developers could understand what the test case is intended for. Further, naming conventions in the test cases followed a similar fashion to the methods they were implemented to test. Moreover, no function dependencies were introduced between the tests and multiple assert statements were avoided as they could lead to confusion where the test failed. Finally, whenever there was a modification in the development, all the tests were run again.

Overall, the use of TDD helped to identify code smells and errors and, thereby, was a noticeable help in increasing the developed code's quality. Further, the ability to repeatedly retest all parts of the application, whenever changes were implemented, increased the confidence in the stability and quality of the code, which is an important factor in a real-world scenario [25].

5 Conclusion

With the increasing data orientation of today's society, the concept of BD is also gaining in importance. However, while its proper use promises immense benefits, ensuring the quality of the corresponding systems is a challenging task. To facilitate this, the application of TDD in the BD domain has been proposed. Therefore, as the underlying research of this paper, a project was carried out to further investigate the general concept. For this purpose, stock market movement prediction was chosen as an exemplary use case and the developed prediction tool and the implementation of the TDD approach were discussed. It was shown that TDD can be useful for BD application development.

Since this approach to Big Data Engineering can be applied to other use cases, extending it to other domains and tools is a promising task for future researchers so that more experience can be gained and collective knowledge can contribute to better TDD process design. Moreover, to further increase the complexity of the current application, an additional overarching prediction service could be implemented that combines the results of the separate prediction algorithms based on the user's preference.

References

1. Staegemann D, Volk M, Jamous N, Turowski K (2020) Exploring the applicability of test driven development in the big data domain. In: Proceedings of the ACIS 2020
2. Shakir A, Staegemann D, Volk M, Jamous N, Turowski K (2021) Towards a concept for building a big data architecture with microservices. In: Proceedings of the 24th international conference on business information systems, pp 83–94
3. Freymann A, Maier F, Schaefer K, Böhnel T (2020) Tackling the Six fundamental challenges of big data in research projects by utilizing a scalable and modular architecture. In: Proceedings of the 5th IoTBDS. SCITEPRESS, pp 249–256
4. Staegemann D, Volk M, Byahatti P, Italiya N, Shantharam S, Chandrashekar A, Turowski K (2022) Implementing test driven development in the big data domain: a movie recommendation system as an exemplary case. In: Proceedings of the 7th IoTBDS. SCITEPRESS, pp 239–248
5. Staegemann D, Volk M, Perera M, Turowski K (2022) Exploring the test driven development of a fraud detection application using the google cloud platform. In: Proceedings of the 14th KMIS. SCITEPRESS, pp 83–94
6. Staegemann D, Volk M, Lautenschlager E, Pohl M, Abdallah M, Turowski K (2021) Applying test driven development in the big data domain—lessons from the literature. In: 2021 International conference on information technology. IEEE, pp. 511–516
7. Crispin L (2006) Driving software quality: how test-driven development impacts software quality. IEEE Softw 23:70–71
8. Shull F, Melnik G, Turhan B, Layman L, Diep M, Erdogmus H (2010) What do we know about test-driven development? IEEE Softw 27:16–19
9. Slaats T, Debois S, Hildebrandt T (2018) Open to change: a theory for iterative test-driven modelling. In: Weske M, Montali M, Weber I, Vom Brocke J (eds) Business process management, vol 11080. Springer International Publishing, Cham, pp 31–47
10. Davies K, Keet CM, Lawrynowicz A (2019) More effective ontology authoring with test-driven development and the TDDonto2 tool. Int J Artif Intell Tools 28

11. Keet CM, Ławrynowicz A (2016) Test-driven development of ontologies. In: Sack H, Blomqvist E, d'Aquin M, Ghidini C, Ponzetto SP, Lange C (eds) The semantic web. Latest advances and new domains, vol 9678. Springer International, pp 642–657
12. Fucci D, Erdogmus H, Turhan B, Oivo M, Juristo N (2017) A dissection of the test-driven development process: does it really matter to test-first or to test-last? IIEEE Trans Softw Eng 43:597–614
13. Beck K (2015) Test-driven development: by example. Addison-Wesley, Boston
14. Williams L, Maximilien EM, Vouk M (2003) Test-driven development as a defect-reduction practice. In: Proceedings of the 14th ISSRE. IEEE, pp 34–45
15. Janzen D, Saiedian H (2005) Test-driven development concepts, taxonomy, and future direction. Computer 38:43–50
16. Sangwan RS, Laplante PA (2006) Test-driven development in large projects. IT Prof 8:25–29
17. Karlesky M, Williams G, Bereza W, Fletcher M (2007) Mocking the embedded world: test-driven development, continuous integration, and design patterns. In: Embedded systems conference. UBM Electronics
18. Shahin M, Ali Babar M, Zhu L (2017) Continuous integration, delivery and deployment: a systematic review on approaches, tools challenges practices. IEEE Access 5:3909–3943
19. Žunić E, Korjenić K, Hodžić K, Đonko D (2020) Application of facebook's prophet algorithm for successful sales forecasting based on Real-world Data. IJCSIT 12:23–36
20. Dash S, Chakraborty C, Giri SK, Pani SK (2021) Intelligent computing on time-series data analysis and prediction of COVID-19 pandemics. Pattern Recogn Lett 151:69–75
21. Ariyo AA, Adewumi AO, Ayo CK (2014) Stock price prediction using the ARIMA model. In: 16th international conference on computer modelling and simulation. IEEE, pp 106–112
22. Bing L, Chan KCC, Ou C (2014) Public sentiment analysis in twitter data for prediction of a company's stock price movements. In: IEEE 11th international conference on e-business engineering. IEEE, pp 232–239
23. Bordino I, Kourtellis N, Laptev N, Billawala Y (2014) Stock trade volume prediction with Yahoo Finance user browsing behavior. In: IEEE 30th international conference on data engineering. IEEE, pp 1168–1173
24. Kolasani SV, Assaf R (2020) Predicting stock movement using sentiment analysis of twitter feed with neural networks. JDAIP 08:309–319
25. Staegemann D, Volk M, Daase C, Turowski K (2020) Discussing relations between dynamic business environments and big data analytics. CSIMQ:58–82

Robust Keystroke Behavior Features for Continuous User Authentication for Online Fraud Detection

Aditya Subash, Insu Song, and Kexin Tao

Abstract Recently, behavioral biometric-based user authentication methods, such as keystroke dynamics, have become a popular alternative to improve security of online platforms, due to their non-invasive nature. However, currently there are very few behavioral biometric authentication methods that provide non-invasive continuous user authentication for online education platforms, resulting in frequent network intrusion and online assessment fraud. Existing approaches mostly analyze the typing behavior of users using a fixed sequence of characters. Furthermore, a better set of features are required to reduce false positive rate for satisfactory performance to prevent online fraud. Existing behavioral analysis methods also mostly rely on conventional machine learning approaches despite recent advancement in deep learning approaches. We identify a set of keystroke behavioral biometric features that yield satisfactory performance by identifying most frequently used features. We also collect new free-form keystroke behavior data during online assessment activities and develop non-invasive continuous authentication methods for free-form text behavior analysis using deep learning approaches. We also compare performance between deep learning and conventional machine learning approaches and evaluate the robustness of the most frequently used features. Result analysis shows that deep learning approaches outperform machine learning approaches on most frequently used feature set. Furthermore, it is found that the identified feature set is robust and results in satisfactory performance in deep learning approaches.

Keywords Online fraud · Most frequently used features · Deep learning · Continuous user authentication · Robustness

A. Subash (✉) · I. Song · K. Tao
College of Science and Engineering, James Cook University, Singapore, Singapore
e-mail: aditya.subash@my.jcu.edu.au

I. Song
e-mail: insu.song@jcu.edu.au

K. Tao
e-mail: kexin.tao@my.jcu.edu.au

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_71

1 Introduction

Recently, research in behavioral biometrics-based authentication systems, especially keystroke dynamics, has gained immense popularity [1] as it is considered a more non-invasive and secure alternative to traditional authentication systems [2, 3]. Keystroke dynamics is not only studied for improving authentication systems, but also other applications including protection of personal identifiable information (PII), thwarting online assessment cheating in online education platforms [4, 5], and digital forensics [6–8]. Furthermore, keystroke dynamics has also made a noticeable impact on the banking, mobile, and health sectors [5].

However, currently there are very few behavioral biometric authentication methods that provide non-invasive continuous user authentication for online education platforms, resulting in frequent network intrusion and online assessment fraud. Existing methods mostly analyze the typing behavior of users using a fixed sequence of characters. Furthermore, a better set of features are required to reduce false positive rate for satisfactory performance to prevent online fraud, which include identity theft and online assessment fraud prevention. Existing behavioral analysis methods also mostly rely on conventional machine learning approaches despite recent advancement in deep learning approaches.

We first identify a feature set that results in satisfactory performance by first reviewing several keystrokes behavioral features. This involves constructing a feature comparison matrix and performing a comprehensive keystroke behavior analysis study to summarize the several keystroke behavioral features, AI learning approaches currently studied, and types of datasets implemented. Furthermore, we collect free-form keystroke behavior data during online assessment activity to investigate and evaluate the effectiveness of identified features for continuous user authentication, for online fraud detection in online education platforms. This includes identity theft and online assessment fraud detection, using sophisticated deep learning approaches. Specifically, we will extract most frequently used features from the feature comparison matrix and conduct detailed evaluation of those features using three deep learning approaches, which include convolution neural network (CNN), recurrent neural network (RNN-LSTM), and transformers.

We also perform a comparative study that compares performance between deep learning and conventional machine learning approaches, which include decision tree (DT), random forest (RF), and k-nearest neighbor (KNN). Lastly, we investigate the robustness of most frequently used features by comparing performance of deep learning approaches trained using the same, most frequently used features from a different dataset.

The main contribution in this paper includes (1) feature comparison matrix for identifying most frequently used features, (2) novel free-form keystroke behavior data collected from India for evaluating most frequently used features for continuous user authentication, for online fraud detection, which include identity theft and online assessment fraud detection in online education platforms, (3) evaluation and effectiveness of most frequently used features using sophisticated deep learning approaches, and (3.1) a sub-contribution to 3, a study that compares performance of deep learning and conventional machine learning approaches and investigates the robustness of most frequently used features.

This paper has been segregated into several sections. Section 2 presents the feature comparison matrix to identify most frequently used features for continuous user authentication for online fraud detection. Section 3 describes the collection of new free-form keystroke behavior data and feature extraction for developing continuous keystroke-based authentication models for continuous authentication, to prevent online fraud. Section 4 presents evaluation of most frequently used features using deep learning approaches and a comparative study that compares their performance with several machine learning approaches. Lastly, an investigation into the robustness of most frequently used features is also performed. We conclude our work in Sect. 5.

2 Keystroke Behavior Analysis

2.1 Feature Comparison Matrix

Keystroke dynamics is rhythmic and temporal patterns generated when a person types. These temporal patterns can be extracted non-invasively from a range of different input devices ranging from normal, virtual, and touch screen keyboards [2, 3, 5]. In this section, we will analyze several research articles and report our analysis. This will include understanding the several categories of datasets implemented, approaches studied, keystroke dynamics research applications, and the features extracted.

According to our analysis, the datasets used in keystroke research include the CMU benchmark [2, 4, 9, 10], Clarkson II uncontrolled free text dataset [11], buffalo partially controlled free text dataset [11], keystroke dynamics Android platform dataset [10], RHU dataset [10], and other novel datasets [6–8, 12–15].

These datasets fall into two categories, namely static (S) [1, 9] and dynamic keystroke (D) data [6, 16]. Static datasets contain keystroke behavior data collected from typing a predetermined string of fixed length multiple times, while dynamic dataset contains free-form keystroke behavior data, which is typed continuously [1, 6, 9, 16]. According to our analysis, majority of studies still rely on static datasets for building AI-based keystroke authentication systems. Numerically, out of 16 studies analyzed, nine of them use static datasets, and eight use dynamic datasets. One of the studies collects both static and dynamics datasets for comparison studies [13].

Datasets consist of several features that have been extracted for training AI learning approaches. The list of which is given in Table 1. This table is called the feature comparison matrix. It is constructed by considering recent studies in information security, which include malware, phishing, intrusion, identity theft, and business email compromise (BEC) detection, with major focus on AI-based methodologies for attack detection. The feature comparison matrix contains summarized information of datasets used, features implemented, approaches studied, their research application, and achieved performance of approaches used in 175 research articles. The feature comparison matrix displayed in this paper is a shortened version of the original matrix containing summarized information of research articles only related to keystroke dynamics. The main objective of this paper is to identify keystroke behavior features for continuous user authentication, for online fraud detection, which include identity theft and online assessment fraud detection, in online education platforms. The construction of feature comparison matrix enabled detailed statistical study that is described in this section. Furthermore, the feature comparison matrix is a major contribution in this paper, as it gives a detailed account of the current trend in the field of keystroke dynamics.

According to our investigation, the most frequently used features fall under the category of digraphs, which includes hold time (H), press press time (DD), and release press time (UD). The feature distribution statistics have been illustrated in Fig. 1. Through Fig. 1 and Table 1, it is noticed that features such as trigraph (T), standard deviation of digraphs (std), average of digraphs (Av), and other features (O) are also implemented, but used less frequently. To extract digraph features, basic data containing key press time (pr) and release times (re) of individual keys are required. Therefore, data collection first involves developing specialized applications, which use certain codes that can retrieve individual pr and re times [6–8]. For example, in

Table 1 Feature comparison matrix

DBAD	ML	DL	Features used									Data used
			H	DD	UU	UD	DU	T	O	Std.	Av	
	✔	✔	✔	✔		✔						S
✔	✔	✔	✔	✔		✔						S
	✔			✔	✔	✔	✔					D
✔			✔			✔			✔			S
	✔	✔		✔	✔	✔	✔					D
✔											✔	D
✔									✔	✔	✔	D
✔		✔	✔	✔		✔						S
	✔	✔	✔	✔							✔	D
✔			✔	✔		✔						S
	✔	✔	✔	✔	✔	✔	✔					D
	✔								✔		✔	D
✔								✔				S
	✔		✔	✔		✔						S
✔	✔	✔	✔	✔		✔						S
	✔		✔	✔		✔		✔				S/D

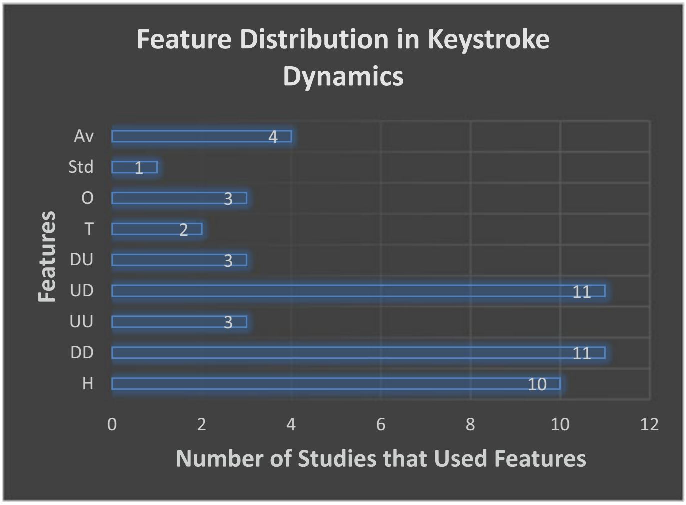


Fig. 1 Distribution of features used in previous keystroke dynamics

JavaScript, EventListeners are used for this purpose. The description of the features, their formulas, and data needed to extract those features have been described in feature calculation matrix (Table 2).

Currently, keystroke behavior research relies on machine learning (ML), deep learning (DL), or distance-based anomaly detectors (DBAD) for user authentication and identification [4, 9, 12]. Our analysis shows that only 50% of research articles analyzed studied DBAD [1, 9, 11, 12, 14], while ~ 44% of them use DL approaches [2, 4, 6–8, 17, 18], and majority studies (~62%) implemented ML approaches [4, 7, 10, 13, 15]. These statistics are inclusive of comparative studies**.**

Research articles also use data preprocessing [13], features selection [7, 10, 11, 13], under sampling [13], optimization techniques [10], and data condensation methods [8]**.** Apart from the research methodology previously mentioned, newer methods of keystroke dynamics analysis have emerged. These include adaptive methodologies that propose to solve evolutionary changes in user behavioral characteristics. It is hypothesized that keystroke behavior changes over time due to several factors such as age and education [4]. According to previous studies, this is solved using proposed real-time behavioral biometric information security system, shortly known as RBBIS system [4]. The proposed system has the capability of building user profiles on the fly and predicting changes in keystroke behavior due to several factors such as age and education overtime, thereby enabling user identification after long periods of time [4].

Overall, we were able to confirm that keystroke dynamics is studied for multiple applications including improving current user authentication and identity verification systems [4, 9, 15], digital forensics [6–8, 11, 13, 16, 19], preventing online assessment

Table 2 Feature calculation matrix

Feature	Description	Formula
Hold time (H)	Time difference between pressing and releasing the same key	pr(i) – re(i), where $i = 0, 1,\ldots,n$, and n is number of keystrokes
Press press time (DD)	Time difference between pressing one key and pressing the next key	pr(i) – pr($i - 1$), where $i = 0, 1,\ldots,n$, and n is number of keystrokes
Release press time (UD)	Time difference between releasing one key and pressing the next key	re(i) – re($i - 1$) where $i = 0, 1,\ldots$,n and n is number of keystrokes
Press release time (DU)	Time difference between releasing one key and pressing the next key	pr(i) – re($i - 1$) where $i = 0, 1,\ldots,n$ and n number of keystrokes
Release release time (UU)	Time difference between releasing one key and releasing the next key	re(i) – re(i–1) where $i = 0, 1,\ldots,n$ and n is number of keystrokes
Trigraph (T)	Latency between alternate keystrokes	Time of key one − Time of key three
Average of digraphs (Av)	Calculated average of all digraphs – D (H, UD, DD, DU, UU times)	$\mathrm{Av} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{D}}{\mathrm{n}}$ where n is number of digraphs
Standard deviation of digraphs (Std)	Calculated standard deviation of all digraphs – D (H, UD, DD, DU, UU times)	$\mathrm{Std} = \sqrt{\frac{\sum_{i=1}^{n} (D-Av)}{n}}$ where n is number of digraphs

fraud in the education sector [4], emotion recognition [13, 16], and adaptive strategies to counter evolutionary changes in keystroke behavior [4].

3 Data Collection

For data collection, two major challenges were considered. The first is related to keystroke data, while the second is related to user profiling. The two questions asked here are “What type of data needs to be collected?” and “how much data needs to be collected from each user for successful user identification?”. We were able answer the first questions using the feature comparison matrix, from which we understand that previous research still relies on static and publicly available datasets for analysis and experimentation. Static datasets contain keystroke behavior data collected from typing a predetermined string of fixed length multiple times [9]. However, this data cannot be used for continuous authentication because users need to be identified at regular intervals of time, and the data received at different intervals of time also varies. Furthermore, there is a lack of publicly available free-form datasets for online fraud detection. Therefore, we decided to collect dynamic or free-form keystroke behavior data for continuous user authentication for online fraud detection. This involves collecting keystroke behavior data from participants during online assessments. We develop our own online education game consisting of several assessment-like activities.

The second question is answered by understanding how previous researchers collected data. Data was collected by requesting participants to type a predetermined password several times, paragraph, or free typing without limitations [8, 9]. These datasets either have multiple session-type records or were collected for long period of time, i.e., 2 years [8, 9]. To collect enough free-form keystroke behavior data and keeping in mind our time constraints, we decided to collect data by requesting participants to play an online education game containing 4 assessment-like keystroke activities, which included 2 answer the question activities and 2 copy the text activities. To get enough volume of data, we requested participants to type at least 100 words per question, because this volume of data received can be used for generating session-type records for continuous user identification.

A total of 13 participants were recruited from Sanjay Gandhi College of Education, Bangalore, India. In addition to requesting keystroke data, participants were asked to fill 2 questionnaires. These questionnaires recorded demographic information such as age, education qualification, handedness, and computer usage statistics. The information collected through these questionnaires will be used for future research applications.

3.1 Preprocessing and Feature Extraction

Raw data received included timestamp, keys pressed, keys released, press time, and release times. The combination of each of these data samples is considered as a record. The raw data was converted into suitable format using Python inbuilt libraries for feature extraction. After the data was received, all keystroke activities for each user were taken separately and merged. Matching records between keys pressed and keys released were extracted due to the presence of repetitive keys. Repetitive keys record several key presses events but record only one key release event, which causes an imbalance between the two events.

To create features, each user's keystrokes are divided into five characters length records, resulting in session like data. This step was performed to simulate session-type data, as the data collected was done only once and not multiple times compared to other publicly available datasets, such as CMU benchmark dataset [9].

The question asked in this subsection is "Which feature set must be selected and extracted for continuous feature authentication, and why?". Due to the presence of various features used for representing keystroke behavior, identifying the right features that yields satisfactory performance is challenging.

Therefore, we propose to use and evaluate the effectiveness of the most frequently used features. According to our analysis, H, DD, and UD times are the most frequently used features compared to other features summarized in the feature comparison matrix. On analysis, we find that H, UD, and DD used as a feature set which yields accuracy of > 90% in machine and deep learning approaches such as CNN, SVM, and MLP [2, 4, 17]. However, this is noticed mainly in studies that use static datasets. Through further analysis, we ascertain that few studies that implement other feature sets for other research applications achieve less accuracy [1, 6–8, 13] compared to the H, UD, and DD feature set. Therefore, we use the most frequently used features, which include H, DD, and UD for continuous user authentication for online fraud detection.

4 Evaluation of Deep Learning Approaches for Keystroke-Based Continuous User Authentication

We propose to use deep learning approaches, namely CNN, RNN, and transformers for keystroke-based continuous user authentication for online fraud detection. To prove that deep learning approaches are more suitable for this application, we compare their performance with three conventional machine learning approaches, which include decision tree (DT), random forest (RF), and k-nearest neighbors (KNN). The architecture of the deep learning approaches is illustrated in Figs. 2, 3, and 4.

To train the approaches, H, UD, and DD digraph features are first extracted from the newly collected free-form keystroke behavior data and split into two sections, namely train (80%) and test (20%). Training epochs were set to 100, while the learning rate was set to 0.001 with cross-entropy loss for multi-class classification. The entire experimentation was performed using Python.

Evaluation of approaches was performed using accuracy (Acc), precision (Pre), recall (Rec), and time to train (TTT). The results are mentioned in Table 3.

According to the result analysis, RNN achieves the highest accuracy, precision, and recall of 89% and 83%, respectively. Specifically, RNN outperforms all deep learning and machine learning approaches. From Table 3, it is evident that CNN achieves the lowest (62%) recall rate among the deep learning approaches. Lastly, transformers achieve the lowest accuracy and precision rate, of 78% and 67% compared to CNN and RNN. It is also evident that deep learning approaches achieve superior performance compared to machine learning approaches in all evaluation criteria. Therefore, from this experiment we can say that deep learning approaches are more suitable for keystroke-based continuous user authentication, for online fraud detection in online education platforms.

4.1 Evaluating Robustness of Features

In the previous section, we were able to confirm that deep learning approaches are more suitable for keystroke-based continuous authentication, for online fraud detection. In this section, we investigate the robustness of features by only training

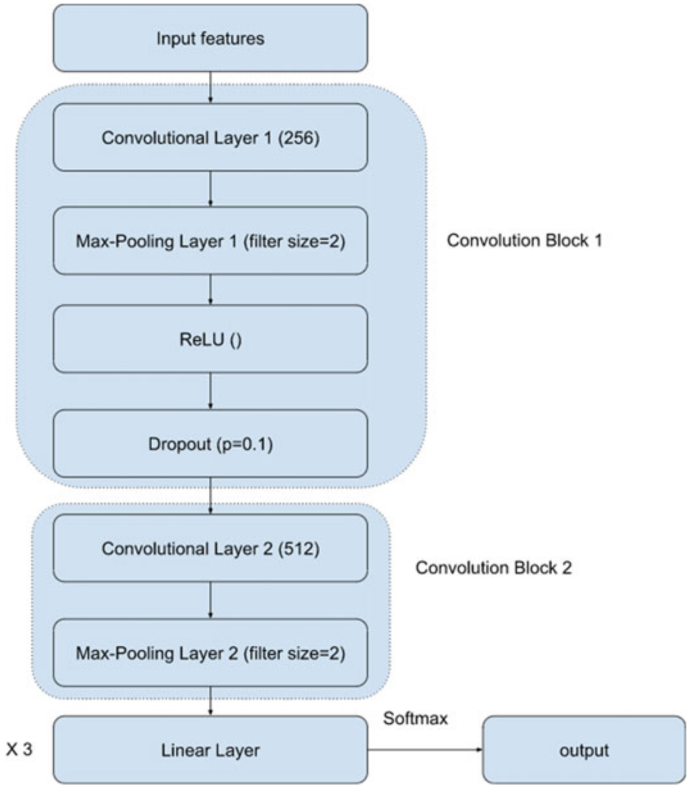


Fig. 2 A1 dimension CNN architecture

the deep learning approaches with a publicly available dataset having same features as the one's identified in Sect. 3.1. We will then compare their performance using the newly collected free-form keystroke behavior data. Specifically, we use the CMU benchmark dataset [9]**.** The dataset contains the same features, to the ones identified and extracted from the new free-form keystroke behavior data collected. This includes H, UD, and DD times. CNN, RNN, and transformers are trained using the same hyperparameters and train test split percentage for result analysis.

According to analysis, RNN achieves an accuracy, precision, and recall rate of 89, 83, and 83 when trained with the new free-form keystroke behavior data and 81, 72, and 71 when trained using CMU benchmark dataset. Despite the collected data having only 10% of the total number of records present in CMU benchmark dataset, the performance achieved using newly collected data is higher for RNN. Numerically, there is an increase of more than 10% in almost all evaluation criteria when new free-form keystroke behavior data is used for training.

However, this trend is not noticed in CNN and transformers. CNN achieves comparable performance in most evaluation criteria, except recall, which has a noticeable difference in value. Transformers achieve comparable performance with very minor

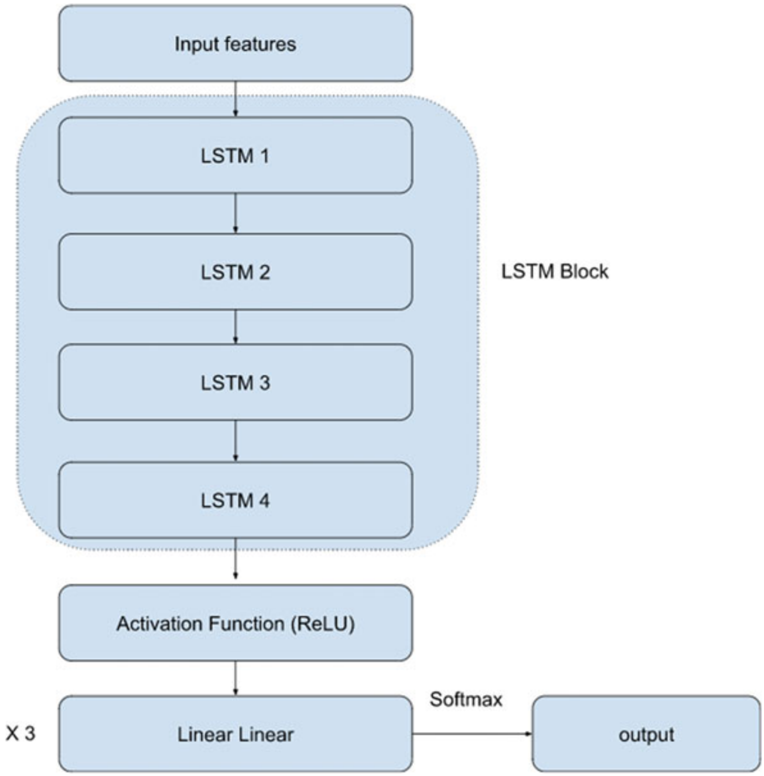


Fig. 3 RNN architecture

difference in accuracy, precision, and recall. The results achieved are illustrated in Tables 3 and 4.

With this experimentation, we were able to investigate the robustness of popular features. According to results achieved, it is evident that performance between the deep learning approaches when trained with different datasets with the same features either have superior or comparable performance in most evaluation criteria, with minor differences in accuracy, precision, and recall. In other words, based on the result analysis the frequently used features that have been identified for continuous user authentication and for online fraud detection are robust and yield satisfactory performance when they were evaluated using deep learning approaches.

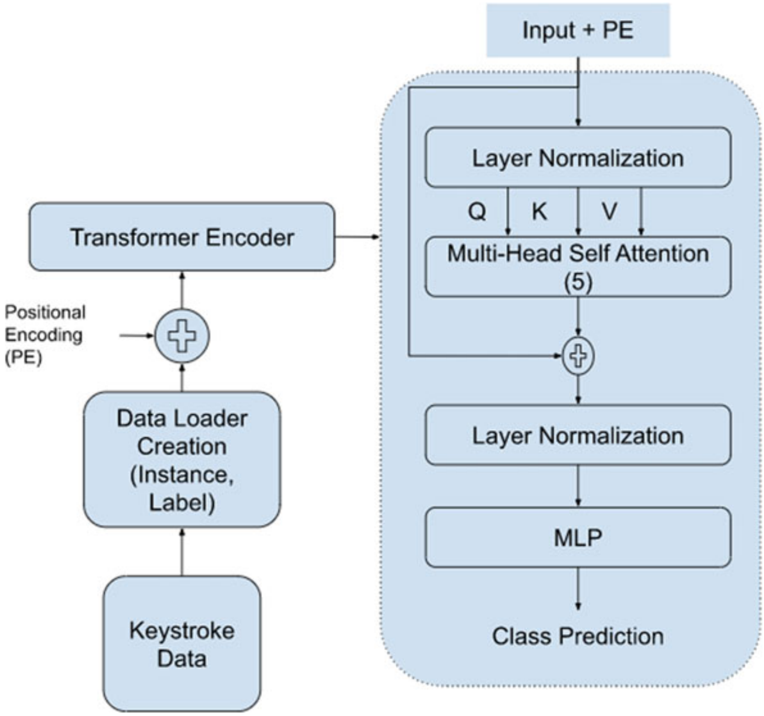


Fig. 4 Transformer architecture

Table 3 Performance comparison of deep learning and machine learning approaches

Algorithm	Acc (%)	Pre (%)	Rec (%)	TTT (mins)
CNN	89	75	62	0.9036
RNN*	89	83	83	0.7009
Transformer	78	67	78	0.2887
DT	48	27	27	0.0007
RF	55	28	28	0.0078
KNN	45	22	22	0.0012

Table 4 Performance of deep learning approaches using CMU dataset

Algorithm	Acc (%)	Pre (%)	Rec (%)	TTT (mins)
CNN	84	80	78	19.30
RNN	81	72	71	3.607
Transformer	79	69	70	2.143

5 Conclusion

In this study, we presented a feature comparison matrix through which we were able to understand and describe the contributions in the field of keystroke dynamics. The feature comparison matrix summarizes the features used, the approaches studied, and the datasets implemented by previous research studies. Using the feature comparison matrix, we were able to identity and evaluate a feature set containing most frequently used features using deep learning approaches. Through experimentation and evaluation, we were able to ascertain that deep learning approaches not only achieve satisfactory performance, but also outperformed machine learning approaches when they were trained with the identified feature set that was extracted from newly collected free-form keystroke behavior data. In other words, the experimentation proved that deep learning approaches are more suitable for keystroke-based continuous user authentication for online fraud detection. Furthermore, we were able to determine that the identified feature set not only yields satisfactory performance but is also robust. This is evident as the performance between the deep learning approaches when trained with different datasets containing the same features either produces superior or comparable performance in most evaluation criteria with minor differences in accuracy, precision, and recall. The contribution presented in this paper will improve current online education platforms by preventing online fraud, including identity theft and online assessment fraud non-invasively. Keystroke dynamics is an important field that is still in its infancy stages, and further research into the field will not only improve security of current authentication system, but also thwart cybersecurity attacks associated with identity theft and improve systems implemented in other sensitive sectors, such as health and online banking.

References

1. Kochegurova EA, Martynova YA (2020) Aspects of continuous user identification based on free texts and hidden monitoring. Program Comput Softw 46(1):12–24. https://doi.org/10.1134/S0361768820010003X
2. Andrean A, Jayabalan M, Thiruchelvam V (2020) Keystroke dynamics based user authentication using deep multilayer perceptron. Int J Mach Learn Comput 10(1):134–139
3. Jain AK, Ross A, Pankanti S (2006) Biometrics: a tool for information security. IEEE Trans Inf Forensics Secur 1(2):125–143. https://doi.org/10.1109/TIFS.2006.873653
4. Subash A, Song I (2021) Real-time behavioral biometric information security system for assessment fraud detection. In: 2021 IEEE international conference on computing (ICOCO), pp 186–191. https://doi.org/10.1109/ICOCO53166.2021.9673568
5. Sadikan SFN, Ramli AA, Fudzee MFM (2019) A survey paper on keystroke dynamics authentication for current applications. AIP Conf Proc 2173(1). https://doi.org/10.1063/1.5133925
6. Tsimperidis I, Rostami S, Katos V (2017) Age detection through keystroke dynamics from user authentication failures. Int J Digital Crime Forensics (IJDCF) 9(1):1–16
7. Tsimperidis I, Arampatzis A, Karakos A (2018) Keystroke dynamics features for gender recognition. Digit Investig 24:4–10. https://doi.org/10.1016/j.diin.2018.01.018

8. Tsimperidis I et al (2020). R 2 BN: an adaptive model for keystroke-dynamics-based educational level classification. IEEE Trans Cybern 50(2)525
9. Killourhy KS, Maxion RA (2009) Comparing anomaly-detection algorithms for keystroke dynamics. In: 2009 IEEE/IFIP international conference on dependable systems & networks, pp 125–134. https://doi.org/10.1109/DSN.2009.5270346
10. Wu T et al (2019) User identification by keystroke dynamics using improved binary particle swarm optimization. Int J Bio-Inspired Comput 14(3):171. https://doi.org/10.1504/ijbic.2019.103613
11. Ayotte B et al (2020) Fast free-text authentication via instance-based keystroke dynamics. IEEE Trans Biometrics, Behavior, Identity Sci 2(4):377–387. https://doi.org/10.1109/TBIOM.2020.3003988
12. Bergadano F, Gunetti D, Picardi C (2002) User authentication through keystroke dynamics. ACM Trans Inf Syst Secur 5(4):367–397. https://doi.org/10.1145/581271.581272
13. Epp C, Lippold M, Mandryk RL (2011) Identifying emotional states using keystroke dynamics. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 715–724. https://doi.org/10.1145/1978942.1979046
14. Bours (2012) Continuous keystroke dynamics: a different perspective towards biometric evaluation. Inf Secur Tech Report 17(1–2):36–43. https://doi.org/10.1016/j.istr.2012.02.001
15. Wu C et al (2018) Keystroke dynamics enabled authentication and identification using triboelectric nanogenerator array. Materials Today (Kidlington, England) 21(3):216–222. https://doi.org/10.1016/j.mattod.2018.01.006
16. Maalej A, Kallel I (2020) Does keystroke dynamics tell us about emotions? A systematic literature review and dataset construction. In: 2020 16th international conference on intelligent environments (IE). IEEE, pp 60–67. https://doi.org/10.1109/IE49459.2020.9155004
17. Maheshwary S, Ganguly S, Pudi V (2017) Deep secure: a fast and simple neural network based approach for user authentication and identification via keystroke dynamics. In: IWAISe: first international workshop on artificial intelligence in security, vol 59
18. Ceker H, Upadhyaya S (2016) Adaptive techniques for intra-user variability in keystroke dynamics. In: 2016 IEEE 8th international conference on biometrics theory, applications and systems (BTAS), pp 1–6. https://doi.org/10.1109/BTAS.2016.7791156
19. Buker RG, Vinciarelli A, Cambria E (2019) Type like a man! inferring gender from keystroke dynamics in live-chats. IEEE Intell Syst 34(6):53–59. https://doi.org/10.1109/MIS.2019.2948514

CPU Benchmarking of the Scalability and Power Consumption of Virtualized Edge Devices

Jeffrey McCann, Sean McGrath, Colin Flanagan, and Xiaoxiao Liu

Abstract Where use cases demand lower latency analysis of streaming video data, the move from traditional cloud-based infrastructure toward computing platforms closer to the edge of the network. Hardware manufacturers are releasing more powerful enterprise-grade server platforms designed to operate in edge environments. Alongside the hardware, the use of virtualization software enables multiple use cases to run concurrently on one physical platform. This paper examines the CPU performance of a physical unit, and the CPU performance of virtual devices, running on the same physical device and demonstrates the benefits virtualization can offer when delivering workloads requiring high CPU workloads, and can also offer benefits in power utilization compared with discrete individual devices.

Keywords Edge · MEC · Virtualization scalability · Video analytics · Edge workload

J. McCann (✉)
Dell Technologies, Limerick, Ireland
e-mail: jeff_mccann@dell.com

S. McGrath · C. Flanagan · X. Liu
University of Limerick, Limerick, Ireland
e-mail: sean.mcgrath@ul.ie

C. Flanagan
e-mail: Colin.Flanagan@ul.ie

X. Liu
e-mail: Xiaoxiao.liu@ul.ie

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_72

1 Introduction

Alam, Ullah [1] discuss the key components of cloud computing that have made it such a success, the primary benefit to users being the ease of deploying and managing workloads. Originally, 'cloud' referred to public cloud datacenters that offered platform as a service (PaaS) and software as a service (SaaS), but the benefits of the 'aaS' model, providing benefits including optimum use of hardware as workloads run on shared, centrally managed hardware platforms. Ease of deployment of new workloads using templates to define the underlying operating system and application layers ensure uniformity of deployment across an organization, thereby improving the security mode, by reducing the number of configuration variables within the deployed systems.

Due to the benefits offered through the cloud model, many organizations have moved to a cloud-style or 'private cloud' deployment model internally within their own datacenters. With the use of converged and hyper-converged platforms such as VMWare Cloud, Amazon Web Services (AWS) Outposts and Snowball systems [2], Microsoft Azure Stack [3] and Redhat Virtualization, which run on Intel x86-based infrastructure, workloads can be deployed into on-premise datacenter environments. A third model, the 'hybrid cloud' model, also exists, where a private cloud model can 'burst' to the public cloud at short notice [4] when extra compute or storage capacity is required.

As workload datasets increase, especially in use cases where computer vision demands high speed, low latency network connectivity, network capacity to transport video streams and process them provide challenges for public cloud platforms, where network latency can account for up to 200 ms per connection. The growth in the use of artificial intelligence (AI) platforms, especially those of neural networks used to undertake object recognition in streaming video, has pushed the requirement for compute platforms away from the cloud and closer to the source of the video data, toward the edge of the network [5]. Sunyaev [6] reviews edge computing and identifies the benefits edge platforms aim to provide in overcoming the challenges posed by processing workloads in the cloud, including reliability and data sovereignty, with the key goal of moving workloads toward the edge of the network to overcome network latency. Reliable network connectivity cannot be guaranteed when wireless or cellular connectivity is used, especially with a computer installed on vehicles.

Computer manufacturers such as Dell Technologies and Hewlett Packard (HP) [3] are beginning to offer enterprise-grade edge hardware platforms designed to be utilized outside traditional datacenter environments. Systems can be designed as a bespoke system for a specific workload or use commercial-off-the-shelf (COTS) hardware platforms, with systems chosen to run a specific workload and utilizing a traditional operating system such as Microsoft Windows or Linux, or more powerful server-based edge platforms, designed to run virtualization software, allowing cloud-like management and deployment features for different workloads on the on-premise devices.

These systems are designed to operate outside of the traditional datacenter environment [7] and so have to offer changes to traditional server design, including form factor, where systems may be deployed in an environment that does not have a datacenter rack, e.g., connected to a machine on a manufacturing line, or short-depth servers that can be installed into existing networking closets, e.g., in retail stores. Enhanced operating temperatures are also required on the systems, with some designed to be sealed units using passive cooling, so debris cannot be ingested into the system in dusty or dirty environments.

This paper benchmarks COTS-based edge devices, a Dell 5200 Edge Technologies Gateway device and a Dell XR11 enterprise grade x86 server with both an ubuntu operating system. The XR11 server was then rebuilt with VMWare ESXi hypervisor, allowing the same performance benchmarks to be run using virtual machines deployed upon ESXi. Power utilization was also captured during the tests to understand each system's performance per watt during the tests.

2 Hardware Equipment Used

2.1 Edge Gateway

Dell 3200 Edge Gateway (shown in Fig. 1) was used as a lower end, industrial grade edge gateway device to provide a baseline hardware platform to compare performance of the virtualized workloads to. The EGW3200 is a COTS device built on the four-core Intel Atom processor with no moving parts, designed to operate in industrial environments with passive cooling and an operating temperature range of –20 to + 60 °C. The system is designed and certified to run either Microsoft Windows or a Linux operating system. System specifications used in the tests are listed in Table 1.

2.2 Server

The Dell XR11 server (shown in Fig. 2) is an enterprise device designed to be operated in a rack environment outside of datacenters, with all cabling at the front of the system to aid access. It has six high-powered internal fan systems to provide airflow across the system, enabling the system to operate at maximum ambient temperatures of 55 °C (dependent on processor/PSU configuration). The XR11 server offers out-of-band (OOB) management capabilities, providing system monitoring, heartbeat monitoring and automated restarting systems where the operating system has crashed. The OOB monitoring platform also allows for remote management of devices outside the system operating through APIs [8] or directly from the Microsoft Azure cloud platform. The specification of the XR11 system is listed in Table 1.

Fig. 1 Dell 3200 edge gateway

Table 1 Hardware specifications

Specification	Dell 3200 EGW	Dell XR11 server
Processor	Intel atom x6425RE	Intel Xeon Gold 6338N
Processor speed	1.90 GHz	3.50 GHz
Core count	4	32
Thread count	4	64
Motherboard	Dell EMC 0d370	Dell 0P2RNT
Chipset	Elkhart lake	Ice lake
Memory	8 GB	128 GB
Disk	240 GB	2 × 1920 GB
File-system	ext4	ext4
Mount options	realtime rw	realtime rw
Block size	4096	4096
Accelerator	–	NVIDIA ampere A2
Operating system	Ubuntu 20.04	Ubuntu 20.04
Kernel	5.13.0-1009-intel (x86_64)	5.4.0-124-generic (x86_64)
Spec sheet	3200 Manual [9]	XR11 Manual [10]
RRSP cost (31/8/2022)	$989	$15,785
Cooling system	Passive	Active
Operating temps	– 20 °C to 60 °C with 0.6 m/s air flow	– 5 °C to 55 °C ASHRAE A2-4 and rugged specs

Fig. 2 Dell XR11 ruggedized server

3 Software Used

3.1 Operating Systems

All systems were installed with a fully patched Ubuntu 20.04.04 server operating system, to provide a consistent OS platform for the tests. Minor kernel build versions for Ubuntu varied due to processor, Intel chipset and VMWare hypervisor configurations.

3.2 Hypervisor

VMWare ESXi Hypervisor version 7.0 Update 3 Build-1864423) was installed to the BOSS card on the XR11, with VMs built and stored on internal NVMe drives. GPU was not enabled within the hypervisor for the purpose of these tests.

3.3 Virtual Machines

Within the hypervisor, three virtual machines (VM) were built using Ubuntu 20.04 server to undertake the performance testing. Table 2 defines the configuration of each system. The configuration for each virtual machine was identical, with the exception of core count and memory, where 4, 8 and 16 cores were provisioned. Each device had 4, 8 or 16 GB of memory assigned within the virtual machine. To ensure the optimum installation of processor components and configuration into the virtual machines, each machine was built from the Ubuntu server install media manually, rather than cloning an existing virtual machine on the hypervisor.

3.4 Benchmark Tools

Each system had a suite of applications installed to manage the system, and to perform the benchmarking tests. These applications are defined in Table 3.

Table 2 Virtual machine specification

Specification	VM1	VM2	VM3
Processor	Intel Xeon Gold 6338N		
Speed	3.5 GHz		
Core count	4	8	16
Thread count	4	8	16
Memory	8 GB	4 GB	16 GB
Disk	240 GB		
File-system	ext4		
Mount options	realtime rw		
Block size	4096		
OS	Ubuntu 20.04 server		
Kernel	5.4.0–124-generic (x86_64)		

Table 3 Application suite installed

Application	Version	Description
p7zip-full	16.02	7-zip file compression software and MIPS benchmark
net-tools	1.60	Assorted network management and reporting tools
gdebi-core	0.9.5.7	Tool to install.deb debian applications (required to install photonix app suite)
Glances	3.2.7	System monitoring too
Docker	20.10.12	Containerization platform

3.5 Power Monitoring System

As used in [11], to capture the power consumption of the system under load, it was necessary to capture the power utilization of the system while at rest and under load. A CT clamp was placed on the power supply to the XR11 server, and the volt-ampere (VA) and power factor (PF) of the device under load were captured at a one-minute interval using an Episensor™ ZEM-61 electricity monitor.[1] The cabling layout of the power monitoring system can be seen in Fig. 3. The ZEM-61 power monitor is connected via Zigbee to a Dell 3000 Edge gateway, which transfers the data via MQTT to an MSSQL database running on a Windows 2019 server.

[1] https://episensor.com/documentation/product-zem-61/.

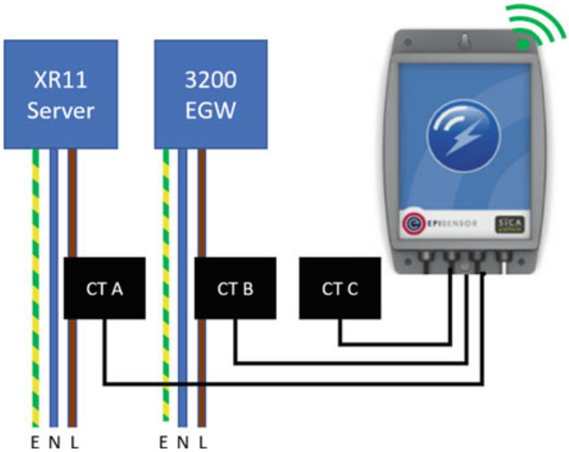


Fig. 3 Power monitoring equipment

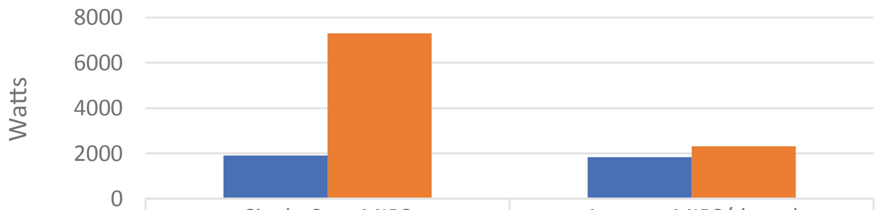
4 Tests

Benchmarking tests were designed which could be ran on each of the physical and virtual platforms. Tests were ran on the 3200EGW and the XR11 with Ubuntu 20.04 installed, and results captured to be used as baseline results. The XR11 was then rebuilt with a VMWare ESXi hypervisor installed, and tests were repeated for each of the individual VMs. The tests were ran on each of the VMs, while they were the only workload running on the hypervisor. VM3 was then cloned and eight identical VMs were then ran concurrently, to test performance difference with multiple workloads running on the hypervisor.

The file compression application 7-Zip provides a benchmarking tool to test processor performance. The application runs four passes of a compression and decompression algorithm on a standardized dataset and uses this to estimate the MIPS performance of a system. MIPs test was run multiple times, and was also run on a single core, and with the maximum number of cores available to the system. MIPs per GHz were extrapolated from the measured results, as well as processor speed. The tests were ran both on a single core, and on MAXCores available on each physical and virtual device.

4.1 Physical System Results

The single thread tests returned a MIPS performance of 1915.5 Mips ($\sum$5.46) for the 3200EGW and 5255.9 MIPS ($\sum$150.89). The test was then repeated with the maximum cores/threads enabled (4 cores/one thread per core vs. 32 cores/ two threads), returning results of 7313.5 Mips ($\sum$20.49) vs. and 147,633.4MIPS ($\sum$1992.91). As the processor in each device was considerably different (1.9 GHz



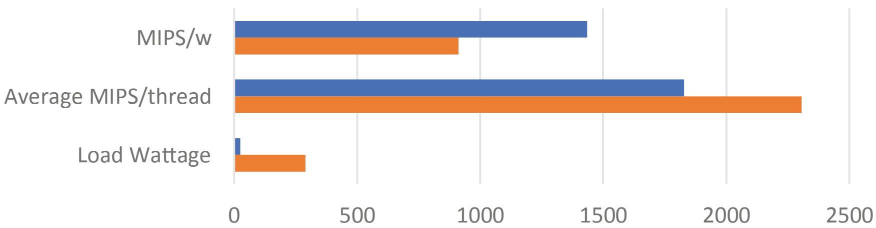
	Single Core MIPS	Average MIPS/thread
EGW3200	1915.5	1828.4
XR11	7313.5	2306.8

Fig. 4 Performance per core/GHz—Physical Machines

vs 3.5 GHz), performance was then extrapolated per GHz, and results are shown in Fig. 4.

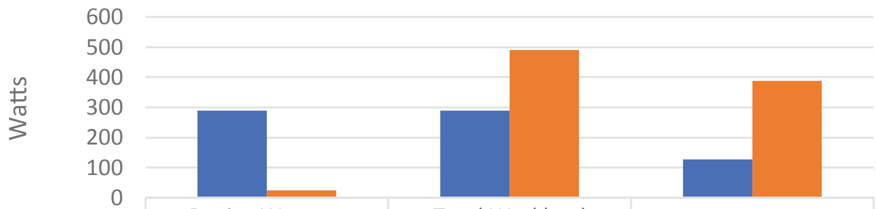
The wattage of the systems was also measured. A baseline wattage was captured from each system at rest (19.4w ∑0.49w vs. 127.8 ∑0.94w) and under load (24.5w ∑1.64w vs. 289.8 ∑27.27w). The average workload (wattage under load-wattage at rest) (6.1w vs 162w) was then used to calculate the MIPs/W of the workload, resulting in MIPS/w of 1434.02 vs. 911.32. Results are shown in Fig. 5.

While the MIPS/W results would show that the EGW3200 provides a performance improvement of 57% for the assigned workload, it is necessary to look at the total wattage that would be required to deliver a total workload of 147,644 MIPS (XR11 performance). To deliver this workload, 21 EGW3200 edge devices would be required. Extrapolating the total wattage requirements to deliver the total MIPs workload on this number of devices is shown in Fig. 6. The total wattage required to deliver the workload on EGW3200s would require 488w to deliver the same performance as the XR11, due to the overheads of underlying baseline system requirements (388 vs 127.8w) between the systems.



	Load Wattage	Average MIPS/thread	MIPS/w
EGW3200	24.5	1828.38	1434.02
XR11	289.8	2306.77	911.32

Fig. 5 Physical device power consumption



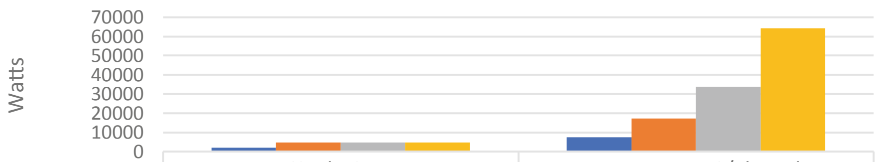
	Device Wattage Under Load	Total Workload Wattage	System Overhead
EGW3200	289.8	289.8	127.8
XR11	24.5	490	388

Fig. 6 Total workload wattage

4.2 Individual Virtualized System Results

The XR11 server was then rebuilt, and baseline wattage was captured of the system at rest. The hypervisor demonstrated a 29.7% increase in power consumption compared with the Ubuntu build (165.79, $\sum$20.03 vs 127.8 $\sum$0.94w) at rest. The tests were then ran sequentially for VM1, VM2 and VM3, with no other load on the system.

Performance across each of the VMs demonstrated similar results for single core (4752.9 $\sum$6.9) and performance of the systems increased linearly with increase in processor and memory configuration when utilizing all cores, VM1 (17,233 $\sum$900.37 MIPS), VM2 (33,975 $\sum$2225.92 MIPS) and VM3 (64,137 $\sum$ 2963.47 MIPS). The MIPs performance results for both single core and average MIPS/thread are shown in Fig. 7. As the systems are all running the same operating system on an Intel x.86 chipset, variations in performance between the 3200EGW and the VMs are directly related to the processor configuration.



	Single Core	Average MIPS/Thread
3200EGW	1915.5	7315.5
VM1	4770	17233
VM2	4773.9	33795
VM3	4714.8	64137

Fig. 7 VM performance results

Comparing the performance (MIPS/(cores * Processor clock speed)) demonstrates up to 7% reduction in performance per core, as core count increased on the VMs (Fig. 8).

Power utilization was also monitored during the tests, with average system averages increasing VM1 (217.22, $\sum$3.22w), VM2 (237.22, $\sum$10.16w) and VM3(258.33, $\sum$79.04w). When removing the underlying average wattage (168.79w) the VMs demonstrate a linear increase in wattage as cores increase (48.54, 71.44, 92.55w).

Figure 9 shows the total power consumption for the workload for each test, with the average MIPS performance per VM, and demonstrates linear increase in power and performance as CPU core count increases.

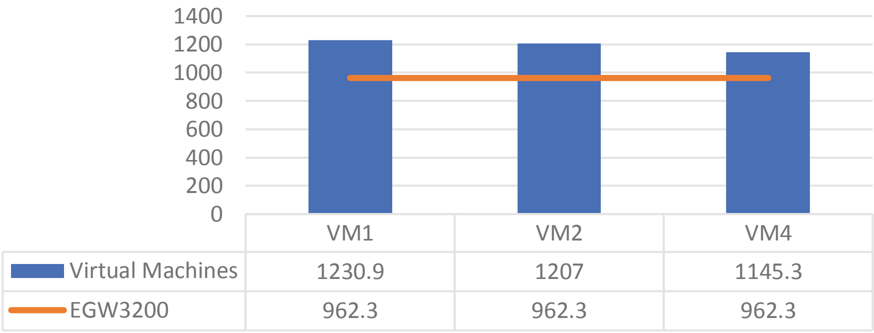


Fig. 8 Performance per core/GHz—Virtual Machines vs 3200

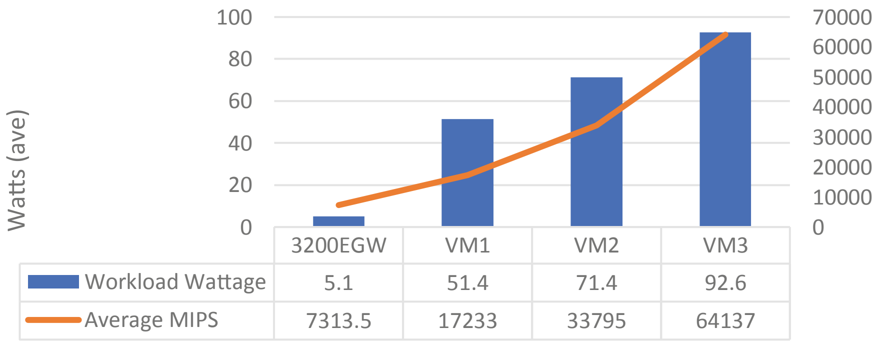


Fig. 9 Power consumption versus performance

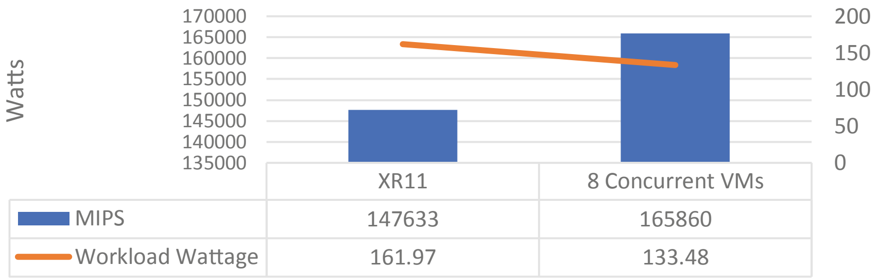


Fig. 10 Concurrent VM performance versus XR11

4.3 *Eight Concurrent Virtual Machines*

The goal and purpose of virtualization is to enable the underlying infrastructure to be fully utilized and run multiple, segregated virtual machines concurrently. To utilize the full underlying server infrastructure (64 cores), eight VM2 (8 GB/8Cores) were cloned, and the tests then ran concurrently. Comparing the performance to the XR11 server running the loads directly, the combined workload (165,860 MIPS) of the eight concurrent virtual machines demonstrated a 12% increase in overall performance compared with the XR11 and a 17.6% decrease in total wattage. (Fig. 10), with 1242.6 MIPS/w average power consumption.

4.4 *Device Cost Versus Performance*

The price for the 3200 EGW and XR11 were retrieved from the Dell website, to allow for cost per MIPs comparison (Table 4). While the XR11 server was almost 16x the purchase price of the EGW3200 when calculating the cost per MIPs, the eight VM virtualized workload provided a 40% reduction in cost per MIPS, compared with the EGW3200.

To provide equivalent compute capabilities using EGW3200 as provided by the XR11 running VMWare ESXi with eight concurrent VMs, 21 3200EGW's costing $20,769 would be required. These costs do not consider the costs of the hypervisor

Table 4 Cost versus performance

	RRSP ($)[a]	Ave MIPS	Cost per MIPs
Dell 3200 EGW	989.00	7314	0.14
Dell XR11 server	15,785.00	147,633	0.11
Eight Concurrent VMs	15,785.00	165,860	0.10

[a] Retail Recommended Selling Prices (RRSP) from dell.com website, 31/9/2022

software, or the cost of installation time for twenty-one 3200EGW systems, vs one XR11 server to provide the same level of compute performance.

5 Conclusion

This research focused on CPU-intensive workloads. The utilization of virtual machines can offer significant performance increase for a specific workload (in this case, 7-Zip Benchmark). A deep understanding of the workload required at the edge, both currently and with projected upcoming workloads, needs to be considered when designing an edge deployment. Where high CPU workloads are identified, the virtualized platforms can also offer benefits in power utilization in comparison with discrete individual devices. The virtualized platform also provides the ability to build virtualized networking capabilities, allowing for segregation and management of the network remotely for management and security purposes.

Alongside the physical performance benefits, other 'soft' benefits including the ability to deploy new workloads to the virtualization platform (network dependent) including: the reduction in onsite visits for deployment of new hardware devices and maintenance of the devices for their lifetime, must be considered during the design of an edge platform.

Next Steps

The research employed off-the-shelf benchmarking tools to understand the performance of general purpose resource-constrained edge devices and the use of enterprise grade edge servers, both using the server as a direct compute device and also as a hypervisor to enable virtualized machines to deliver workloads. The focus for future work following this research, is to how to correctly identify and scale the design of edge devices in the most cost-effective manner for specific use cases, both for now, and for projected new workloads. There are opportunities for further research into quantification of the use of virtualization of edge workloads to provide cost, performance, security and soft benefits in comparison with physical edge devices for differing usecases, or multiple instances of the same usecase.

Areas of Focus

1. The predominant areas of focus for further investigation include:
2. Identification of algorithms used in calculating compute speed and cost in processor utilization;
3. GPU usecase requirements, and ability to share accelerator technologies in virtualized platforms;
4. Management of remote edge platforms;
5. Containerization of code and how this could be utilized to move the compute requirements to different platforms, dependent upon current and projected load;

6. How to handle requests to enable requests for increased compute on edge devices, and the ability to 'burst' to a cloud platform;
7. The goals for ongoing research would be to identify an algorithm or suite of algorithms that can decide in real time, as to the processing point within the technical ecosystem, that meets all of the demands for the resultant information, in a timely and cost-effective manner.

References

1. Alam A, Ullah I, Lee Y-K (2020) Video big data analytics in the cloud: a reference architecture, survey, opportunities, and open research issues. IEEE Access 8:152377–152422
2. Deb M, Choudhury A (2021) Hybrid cloud: a new paradigm in cloud computing. Mach Learn Tech Analytics Cloud Secur:1–23
3. Chawla H, Kathuria H (2019) Building microservices applications on azure stack. Building microservices applications on microsoft azure. Springer, pp 245–255
4. Dreibholz T et al (2019) Mobile edge as part of the multi-cloud ecosystem: a performance study. In: 2019 27th euromicro international conference on parallel, distributed and network-based processing (PDP). IEEE
5. Véstias M (2020) Processing systems for deep learning inference on edge devices. Convergence of artificial intelligence and the internet of things. Springer, pp 213–240
6. Sunyaev A (2020) Fog and edge computing. Internet computing. Springer, pp 237–264
7. Rohith M, Sunil A (2021) Comparative analysis of edge computing and edge devices: key technology in IoT and computer vision applications. In: 2021 international conference on recent trends on electronics, information, communication & technology (RTEICT). IEEE
8. Faasse S, Bucek J, Schmidt D (2020) Out of band performance monitoring of server workloads: leveraging RESTful API to monitor compute resource utilization and performance related metrics for server performance analysis. In: Proceedings of the ACM/SPEC international conference on performance engineering
9. Dell Technologies. Dell Technologies edge gateway 5200 specsheet. 2022 [cited 2022 24/1]; Available from: https://www.delltechnologies.com/asset/en-us/solutions/business-solutions/technical-support/dell-technologies-edge-gateway-5200-spec-sheet.pdf
10. Dell Technologies. *Dell EMC PowerEdge XR11 Technical Specifications*. 2022 [cited 2022 14/10]; Available from: https://dl.dell.com/content/manual23192522-dell-emc-poweredge-xr11-technical-specifications.pdf?language=en-us&ps=true
11. McCann J et al (2022) Benchmarking the scalability & power consumption of general-purpose resource-constrained Edge devices for video stream analysis. Int J Comput Appl 29(3):138–149

Implementation of a Mobile Application for Checking Medicines and Pills for the Visually Impaired in Korea

Soeun Kim, Youngeun Wi, and Jongwoo Lee

Abstract According to the Pharmaceutical Affairs Act of South Korea, medicines should be used correctly and safely by stating the product name, expiration date, and dosage on the container or packaging. However, such matters are rarely marked in braille, and the labeling is insufficient, making it difficult for visually impaired people to accurately know and take medication information. In this paper, we implement a mobile application that provides integrated services, so blind people can conveniently inquire about medicines based on voice guides. Medicines' search is possible through various methods such as container recognition, pill recognition, speech recognition, prescription recognition, and drug envelope recognition. A CNN-based pill recognition model was also developed. Using our implementation, blind people can quickly obtain information about medicine through simple UI and minimal input. As a result, our application will improve access to medicine information for the visually impaired while also reducing misuse by the visually impaired.

Keywords Search for medication · Visual impairments · Pill recognition

1 Introduction

Blind people are limited to information that can be obtained by sight, so they get data using touch and hearing, which are sensory information other than sight. Medicines have a direct or indirect effect on health and life, so we should know and take them correctly, but there are many similar things such as regular medicine packaging, prescribed medicine bags, and pills, making it difficult for blind people to distinguish them just by touching the shape.

S. Kim (✉) · Y. Wi · J. Lee
Sookmyung Women's University, Seoul, Republic of Korea
e-mail: soeunkk99@gmail.com

J. Lee
e-mail: bigrain@sookmyung.ac.kr

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_73

In the case of medicine containers, the Korean government recommends that the main contents of medicines can be marked in braille, but most drugs are not marked in braille. The reason is that braille labeling is just a recommendation, not an obligation, and the actual cost burden of pharmaceutical companies due to packaging changes occurs. According to the Ministry of Food and Drug Safety, only 0.2% of all medicines in 2020 were marked with braille, and even braille-marked medicines have different braille specifications, labeling items, and locations, making braille marking less effective [1].

Blind people are restricted not only in distinguishing medicines but also in information on taking medicines. According to a previous study [1] investigating whether to use medication guidance for the visually impaired, it is not easy to receive medication guidance from a pharmacist, and visually impaired people have no choice but to rely on oral medication guidance. Therefore, there is a risk of misuse of medications if they cannot remember the oral medication guidance even after receiving medication guidance from experts [2].

Therefore, visually impaired people should use other means to inquire about drug information. In this paper, we focus on the applications of these means. Existing services that provide drug search functions have three problems that make it difficult for visually impaired people to use them. Since the detailed analysis is covered in the next chapter, we will list only the problems here. First, information input occurs too frequently. Second, the layout structure is complicated because it is not a service only for the visually impaired. Third, there is a limited way to inquire about medicines.

In this paper, to overcome these problems, we implement a “Pillaroid" application that provides information about medicine to the visually impaired in various ways such as pill recognition, container recognition, speech recognition, prescription recognition, and medicine envelope recognition.

2 Related Research

This section introduces applications that provide existing drug search functions. It also analyzes the disadvantages related to the inconvenience users’ feel in using the existing drug search functions.

2.1 Siloam Healthmore

Siloam Healthmore [3] is a Korean mobile application that provides a drug information search service for the visually impaired. In addition to simple text input, there is a feature that blind people can search for drug information by scanning the barcode and QR code of drugs and searching by voice using built-in microphones. However, Siloam Healthmore does not provide any voice guidance for taking photos for drug searches. Therefore, since it is difficult for the visually impaired to grasp the angle and location of the drug shooting, the probability of inquiring about the drug is significantly reduced when using the actual application, making it inconvenient to inquire about the information. In addition, although visually impaired people are the main target of the service, the number of touches and inputs for service use is too high, and information accessibility is poor by outputting more than three lines of text without a voice guide.

Therefore, Pillaroid proposed in this paper reduces the probability that the drug is not photographed by providing a real-time guide according to the position of the hand when searching for the drug by shooting. In addition, it is possible to search with minimized input by automatically providing voice guides such as function names and descriptions of the screen.

2.2 Drug Search

Drug Search [4] is a Korean application that provides drug information services at the Korea Pharmaceutical Information Center. You can inquire not only by entering the product name or related information but also by entering or selecting the shape of the formulation and color. However, since it is not a service for the visually impaired as a target layer, it is difficult to inquire about information because it has a rather complicated layout arrangement. In addition, it is inconvenient because only text input is possible as a method of searching for drugs. Further, only drug information can be viewed in the application, and there is no personalized service such as favorites and notifications, so the schedule and information must be managed by using additional applications. Therefore, in the Pillaroid, information can be inquired using methods such as shooting and voice recognition, and notification and favorite service functions can be accessed by pressing a single button on the screen that displays the searched drug information.

2.3 ConnectDI

ConnectDI [5] is a Korean mobile application that can search for drug information and offers counseling services with pharmacists through non-face-to-face chat counseling. ConnectDI is characterized by being able to search for information on a drug by entering the name or symptoms of the drug in one text input field, so it can be searched with the minimum number of inputs. In addition, since the drug of interest can be selected, there is an advantage that everyone can be divided into two categories of drugs and injections and inquired individually. However, on the main screen, there is too much information on one screen, such as drug search, news, and channels, making it less efficient for blind people to find the desired function at once. Therefore, Pillaroid aims to increase efficiency and convenience by placing up to two buttons horizontally or vertically in the layout of the menu.

3 System Design

3.1 System Goals

The Pillaroid, a medicine search mobile application for the visually impaired, aims to improve information accessibility for the visually impaired and provide a convenient way to use it. The detailed directions for this are as follows. First, all guidance's of Pillaroid are provided by voice. It provides voice guidance according to the situation during shooting or voice recognition and divides the click method into two so that the information of the text and button can be provided by voice to the user. The user can check the information of the text or button by voice by clicking the component once and can perform the original selection function by double-clicking. Second, information input is minimized in all input processes of the Pillaroid. In particular, the five search functions, which are the main functions, do not go through a complicated input process by recognizing only the shooting or drug names by voice. Finally, the service is simplified by excluding unnecessary functions. To this end, the rest of the functions, except for the favorites and dosage notification functions, were made available without logging in.

3.2 System Diagram

The overall system configuration of the Pillaroid is shown in Fig. 1.

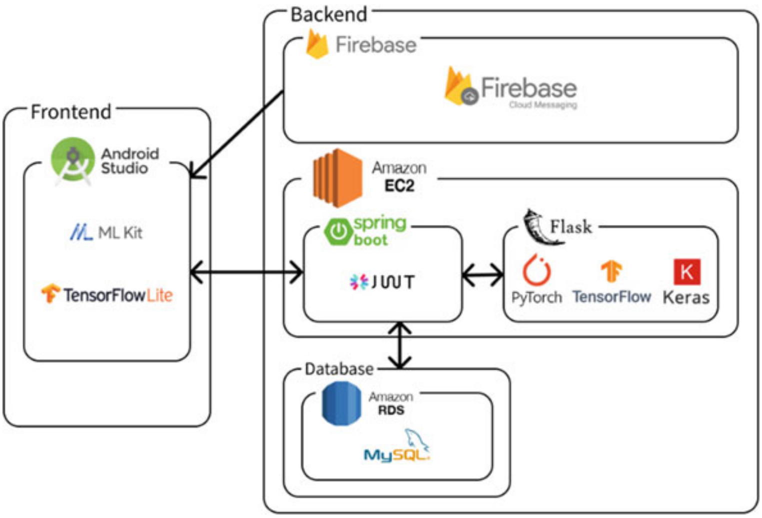


Fig. 1 Structure of the proposed system

The application uses Text-to-Speech (TTS) technology [6] to voice text to blind people. TensorFlow Lite [7] is used to detect a hand in real time when taking a picture of a pill, and the Google ML kit library [8] is used to recognize text and barcodes in pictures when taking containers, prescriptions, and drug envelopes. In addition, the application provides a push notification function by calculating the time to be taken based on the medicine that sets the notification of taking and the mealtime set by the user. To this end, Firebase Cloud Messaging (FCM) [9] provided by Firebase is used.

The application's basic server is implemented with Spring Boot [10]. This server is that communicates directly with the mobile application and handles the client's request by interworking with the database built using MySQL. It handles all other requests except for pill retrieval, and in case of a request for pill retrieval, it sends an image to the deep learning server to request pill identification. The Pillaroid is managed by Kakao Login, and JSON Web Tokens (JWT) [11] was used at this time.

The deep learning server was built using Flask, a Python web framework. Through YOLOv5, a PyTorch object detection model, it checks whether a pill exists in the photo sent to the server and cuts the photo leaving only the detected pill part. In this paper, about 20,000 pill images were additionally trained on the existing YOLOv5 model to specialize in pill detection. To classify pills, we construct a CNN training model based on TensorFlow and Keras, and we obtain predictive results from the model with cropped pill images.

4 Performance Evaluation

As shown in Fig. 2a, the main screen shows four buttons. Among them, you can choose how to search by packaging container or pill in "Searching by Shooting," and "Searching in Document" supports how to search by prescription or medicine envelope.

4.1 Basic Functions

Searching for medications by container recognition At the start of the function, the rear camera is turned on to take a picture of the medicine packaging container, and the user is notified by voice that the camera is turned on. When a user brings a medicine packaging container to the rear camera, Pillaroid attempts to detect barcodes in real-time images through the Google ML kit library (Fig. 2b). If the barcode is not detected, the TTS function will say, "Barcode is not recognized. Please slowly move up, down, left, and right," and then try to detect the barcode again. If the barcode detection is successful, it passes the barcode number to the server to obtain related medicine information. As shown in Fig. 2c, the medicine information is divided into efficacy, usage and dosage, precautions, appearance, ingredient, and storage method.

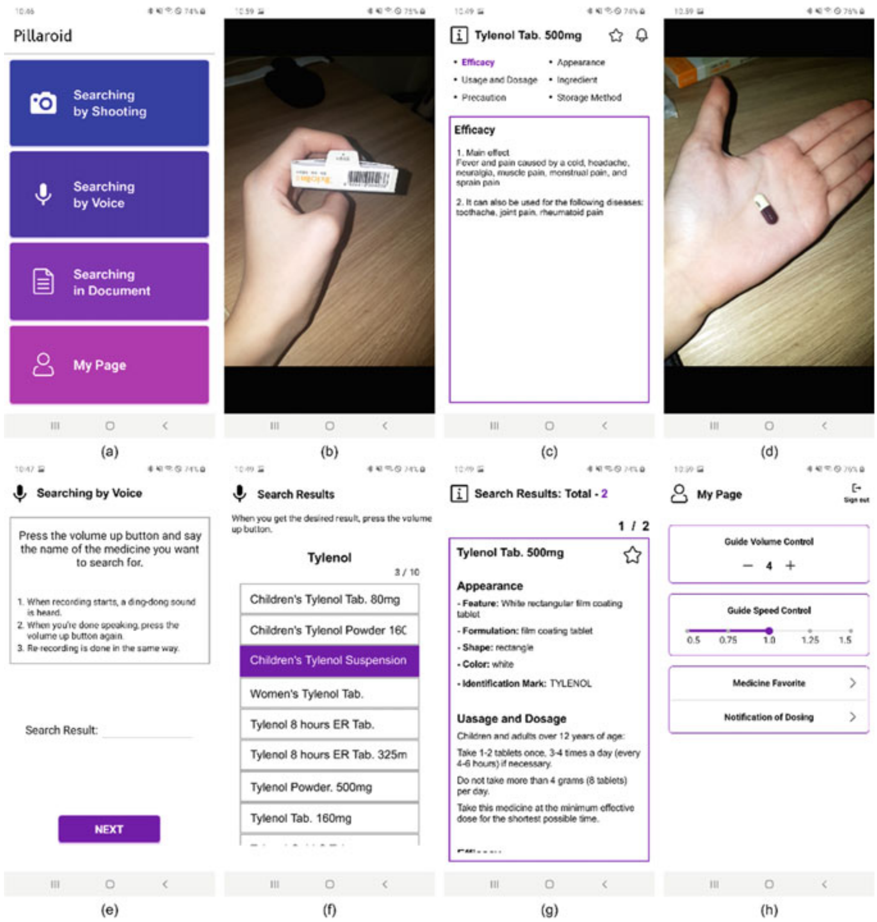


Fig. 2 Screenshots of the proposed Pillaroid application: **a** main, **b** searching by container, **c** medicine information result, **d** searching by pill, **e** searching by voice, **f** list of search results, **g** prescription search results, **h** my page

In the server, the processing is divided according to whether the received barcode number is in the database. If barcode information exists in the database, it finds related drug information and returns it directly to the client. If not, it crawls in "Drug Safety Country" [12] and checks whether there is a medicine with the corresponding barcode number. The serial number and product name are extracted from the drug page found by crawling, and if it is a medicine in the database, the searched information is returned to the client.

Searching for medications by voice In the pharmaceutical voice search screen shown in Fig. 2c, the voice guide for the search method is first automatically executed using Android's TTS function. After that, voice recognition is operated using a volume button. Pressing the volume-up button starts voice recognition with a ding-dong sound. After that, when the user says the name of the drug and presses the volume button again, voice recognition ends. If you press the volume-down button while voice recognition is in progress, voice recognition is stopped. In addition, re-recording proceeds in the same way by pressing the volume-up button again after voice recognition is completed. At the bottom of the screen, there is a confirmation button. The location of the button prevents the user from finding the button by guiding the user by voice to “Button. Check the result” when the button is clicked once. Therefore, when the user finds the location of the button and clicks it after the end of voice recognition, the voice recognition keyword is output to TTS, and the drug name is delivered to the server.

The returned drug names are shown in the form of a list on the Android screen (Fig. 2f). At this time, the list of searched drug names is automatically outputted one by one as a voice. When selecting one of the searched results, the user can double-click the part on the screen or press the button above the volume when the corresponding drug is heard during voice output. Thereafter, the information output screen is the same as the result screen (Fig. 2c) shown when searching by container shooting.

Searching for medications by prescription Pillaroid also supports inquiry of drug information through prescription shots. The text written on the prescription is recognized by the ML kit after being photographed with a camera. Considering the general format of Korean prescriptions, the recognized drug name is extracted between the position where the “name of the drug” is written and the position of the “injection prescription details” or “margin” text. The names are sent to the server when one or more drug names have been extracted. The server searches for drugs that match the term or that begin with it. At this time, if there are multiple results in a single drug name, the drug that exactly matches the drug name except parentheses is given priority. It then returns the fastest drug information of the prior sequence as a response result. The returned drug information is the drug name, appearance information, usage and dosage, efficacy, and effectiveness. The information is sequentially displayed on the screen as shown in Fig. 2g while guiding the drug name search on the application screen to a voice guide. The information for each category is output as a voice when each category area is clicked. When there are more than two results for a drug, it is also possible to check the information for the following drug by enabling horizontal swiping on the screen’s result output portion.

Searching for medications by medicine envelope When the visually impaired want to check whether the medicine they want to take is correct, they can check the information by taking a picture of the medicine envelope. When the rear camera is turned on, the visually impaired can take a picture of the medicine envelope through the volume-up button. The ML Kit text recognition library recognizes the text written on the medicine envelope, where the medicine envelope refers to two types: a sachet and a pharmacy envelope. The name of the pharmacy and the timing of taking it, which is mainly indicated as breakfast, lunch, and dinner, are extracted from the sachet. The pharmacy name, date of manufacture, drug classification, and voice-eye code are extracted from the pharmacy envelope. The voice-eye code is a rectangular code written on a pharmacy envelope, and when you scan it in the "Voice Eye" application [13], you can check the medication information and medicine information by voice. The voice-eye code is determined by the recognition of the text of the "voice medication map" written under the code. The extracted text information is synthesized and provided as a TTS voice guide.

4.2 Search Function Through Pill Shooting

Pill identification using deep learning The Ministry of Food and Drug Safety's open dataset [14] and a dataset collected directly from different lighting, angles, and distances make up the pill learning dataset used in this paper. The types of pills in the open dataset are very vast, so drugs that users are interested in were selected through the degree of user response in Naver's drug dictionary [15]. Through the data augmentation process with brightness adjustment, rotation, and inversion techniques for each selected image, a dataset of about 120,000 images of 3091 types was finally built.

In this paper, CNN was used as a prediction model for pill deep learning. The CNN layer consisted of one input layer, seven hidden layers, and one output layer, and each of the seven hidden layers consisted of four convolutional layers and three pooling layers. The convolution layer uses ReLU as an activation function, and the output layer uses the Softmax function for multi-class classification.

Table 1 shows the results of evaluating the learning model with 100 pictures of pills. The evaluation photo consisted of some of the open datasets that were not used for learning and photos taken by the authors who saw the pills on their hands. Out of a total of 100 photos, there were nine cases where the pill was incorrectly predicted, and the overall accuracy was 91%.

On the other hand, the predicted probability for each pill class makes it possible to estimate uncertainty. For example, the fact that one probability value is significantly

Table 1 Predictive results based on predicted value criteria

Prediction probability (%)	Range (%)	Prediction success	Prediction failure	Sum
100	–	91	9	100
75	More than 75	84	6	90
	Less than 75	7	3	10
70	More than 70	88	6	94
	Less than 70	3	3	6
65	More than 65	88	7	95
	Less than 65	3	2	5

higher than another implies that the model is confident of the result. Conversely, similar probability values mean that the uncertainty is high, and the results are unreliable. Therefore, if a threshold value is set and the highest predicted probability is lower than the threshold value, the judgment should be excluded because it is not reliable. Table 1 shows the results of classifying the success of the prediction based on 75%, 70%, and 65% to determine the optimal threshold. For the predicted value to be optimal, it should be a value that detects fewer prediction failures but more prediction successes. When the prediction probability was 75%, it detected four fewer successful predictions than when it was 70%, and when it was 65%, it detected one more failed prediction than when it was 70%. Therefore, a threshold value of 70% is appropriate, and if the largest predicted probability does not exceed 70%, the predicted result is not returned.

Execution process When you start the pill search function, it notifies you that the rear camera is turned on and sends a voice guide telling you to put the pill on your palm. The app performs analysis through TensorFlow Lite to detect hands in real-time images being filmed with a camera (Fig. 2d). If the hand is not detected at the appropriate distance and position, "Hands not detected. Please move the camera farther away to capture the hand." After providing the guide, it will proceed with the analysis again. If the hand is detected well in the right place, it sends the shot to the Spring Boot server and displays the pill information recognized by the server in the app. The pill information screen is the same as the result screen (Fig. 2c) shown when searching by container photographing. The Spring Boot server sends a picture received with the request to the Flask server to obtain a list of medicine serial numbers. It finds medicine information in the database and returns it to the app with the serial number of the pill that corresponds to the highest prediction.

In the Flask server, the pill prediction data are obtained through the model after checking the existence of the pill in the requested image. YOLOv5 was used to assess whether a pill was present in the picture. Using labelImg [16], an image labeling tool, a set of pill data was converted into learning data to be used in YOLOv5. The pill image recognized by YOLOv5 is sent into the CNN model, which then runs the model to get the pill prediction result. The serial number and predicted value of the top five items with a high probability of prediction are returned to the Spring Boot server.

4.3 Other Function

Pillaroid guides how to use the service, a shooting guide, and text displayed on the screen by voice. At the time of initial execution of the Pillaroid, an initial screen for guiding the app usage description is displayed, and a voice guide is also executed at the same time. After that, when the app was relaunched, the initial guide screen was not visible. In addition, when all elements such as text, images, and buttons displayed on the app are touched once, text or related explanations are guided by voice first.

5 Conclusion

In this paper, we implemented a Pillaroid that provides a method of shooting and voice recognition so that visually impaired people can search for drug information efficiently and conveniently. Since it mainly targets the visually impaired, unlike other applications that provide existing drug information search services, the number of information inputs is minimized, and the layout is simply placed so that only one piece of information can be contained on one screen. In addition, the drug was implemented so that information could be inquired using voice and shooting methods as well as text input.

The main functions of the Pillaroid are to search for drugs by taking a medicine container or individual shots, to check drug information by voice recognition, to check the list of drugs through prescription shots, and to check the contents on the envelope by taking a medicine bag. In this way, the user may get information on the drug by voice or additionally inquire about information. When shooting the container or individual of a drug or searching for a drug by voice recognition, six types of information are output: efficacy, usage, precautions, appearance, ingredients, and storage methods. On the other hand, when searching by prescription, three types of information are shown: appearance, usage, and efficacy.

In the case of a drug search by individual shooting, the user's hand is recognized during the shooting, and information about it is provided as a voice guide. By implementing a hand recognition model, a situation in which the user does not have a pill when taking a pill was minimized. Furthermore, if the hand is recognized at an appropriate location and distance using the model, the medicine information corresponding to the pill is inquired with the screen capture image and displayed on the screen.

When searching for drugs by shooting pills, the user's hand is recognized during the shooting, and information about it is provided as a voice guide. By implementing a hand recognition model, a situation in which the user does not have a pill when taking a pill was minimized. In addition, if the hand is recognized at an appropriate location and distance using the model, the medicine information corresponding to the pill is inquired with the screen capture image and displayed on the screen.

It is obvious that the drug information search function provided by Pillaroid provides convenience to the blind and improves information accessibility. However, because the accuracy of the pill recognition model does not super exceed 90 percent, there is a limitation that the risk of users' misuse of drugs is not zero. We will later develop into a more usable service by modifying and improving the pill recognition model, such as recognizing fine text written on the pill.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2022R1F1A1063408). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF2022H1D8A303739411). This research was supported by the MSIT

(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP (Institute of Information & communications Technology Planning & Evaluation) (2022-0-01087).

References

1. Lee BH, Lee YJ (2019) Evaluation of medication use and pharmacy services for visually impaired persons: perspectives from both visually impaired and community pharmacists. Disabil Health J 12:79–86. https://doi.org/10.1016/j.dhjo.2018.07.012
2. Kim H, Koo H, Oh JM, Han E (2017) Qualitative study for medication use among the hearing impaired in Korea. Korean J Clin Pharm 27(3):178–185. https://doi.org/10.24304/kjcp.2017.27.3.178
3. Google Play. Siloam Center for the blind. https://play.google.com/store/apps/details?id=com.siloam.healthmore&hl=en_US&gl=US
4. Google Play. Korea pharmaceutical information center. https://play.google.com/store/apps/details?id=kr.health.dikmobile&hl=en_US&gl=US
5. Google Play. Connectdi. https://play.google.com/store/apps/details?id=com.connectdi.onesglobal&hl=en_US
6. Text to Speech. https://cloud.google.com/text-to-speech?hl=en
7. Tensorflow Lite. https://www.tensorflow.org/lite?hl=en
8. ML Kit. https://developers.google.com/ml-kit?hl=en
9. Firebase Cloud Messaging. https://firebase.google.com/docs/cloud-messaging?hl=en
10. Spring Boot. https://spring.io/projects/spring-boot
11. JWT. https://jwt.io/
12. Nedrug. https://nedrug.mfds.go.kr/index
13. Google Play. Voiceye. https://play.google.com/store/apps/details?id=com.voiceye.reader.access&hl=en_US
14. Pill Identification Dataset. https://nedrug.mfds.go.kr/pbp/CCBGA01/getItem?totalPages=4&limit=10&page=2&&openDataInfoSeq=11&hl=en#none
15. Naver Pharmaceutical Dictionary. https://terms.naver.com/medicineSearch.naver
16. GitHub. labelImg. https://github.com/heartexlabs/labelImg

Air Traffic Management System Business Process Analysis for the Development of Information Exchange Interoperability Framework

Anwar Awang Man, Ab Razak Che Hussin, and Okfalisa Saktioto

Abstract Air traffic management (ATM) system is one of the tools that is used to manage and control air traffic by providing air traffic controllers the information needed to make effective and safe decisions in their daily operations. The source of data for ATM system comes from various sub-system within the air traffic management ecosystem. These data are fused together to form valuable information for use by the air traffic controllers in making accurate and safe decisions. As the volume of air traffic increases, the current method of exchanging data has become a challenge to the interoperability between ATM system and its sources of data. A modern and end effective way needs to be established to address this issue. Understanding of each business process for each sub-system that contribute these data is important to identify the interoperability issues and challenges in exchanging these data. This paper focuses to identify the current business process involved in ATM system information exchanges within the Civil Aviation Authority of Malaysia (CAAM) through brainstorming method. Findings from the brainstorming session will be documented using mind mapping method and the identified business processes involved will be documented by using the BPMN 2.0 notation. The findings will then be further use to develop a usable and logical information exchange interoperability framework for CAAM ATM System.

Keywords Air traffic management (ATM) · ATM system · Information exchange · Interoperability · Business process · ASBU · SWIM

A. Awang Man (✉) · A. R. Che Hussin
Faculty of Management, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia
e-mail: anwar73@graduate.utm.my

A. R. Che Hussin
e-mail: abrazak@utm.my

O. Saktioto
Informatics Engineering Faculty, Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim, Riau, Indonesia
e-mail: okfalisa@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_74

1 Introduction

Air traffic management (ATM) systems face significant challenges because of the demand for enhanced safety, efficiency, and capacity resulting from the rapid expansion of global air transport and the growing concern for environmental sustainability issues. In this perspective, the International Civil Aviation Organization's (ICAO) Global Air Navigation Capacity and Efficiency Plan (GANP) identifies the following critical performance improvement areas:

- Airport operations.
- Efficient flight path planning and execution.
- Optimum capacity and flexible flights.
- Globally interoperable systems and data.

Within the next two decades, the air transportation sector is likely to develop dramatically. Clearly, this growth might have a severe influence on the environment if the sustainability concerns are not addressed [1]. Innovative ATM and avionics systems can have an immediate influence on mitigating aviation's environmental consequences, while several eco-friendly technical solutions are being considered for tackling the long-term issues posed by the aviation industry's constant expansion. Several international and regional research efforts are currently tackling ATM modernization concerns [2]. However, these efforts and programs usually offer unique solution toward certain region or states and require some degree of customization to be adopted by other state.

This paper will examine and understand the current business process of ATM with focus on the improvement of global system interoperability which is directly related to the performance of ATM systems within the Civil Aviation Authority of Malaysia (CAAM). Findings from this analysis shall be the basis of improvement action on the current process and shall be further use as a basis to formulate an ATM System Information Exchange Interoperability Framework for CAAM.

2 What Is Air Traffic Management (ATM)?

2.1 Definition of ATM

Clear understanding of ATM definition is the first step of understanding the business process. The term air traffic management (ATM) can be defined as a discipline of managing aviation traffic and its related resources such as air space and flight route. ATM services includes air traffic control services, flight route management, and air traffic flow management which its objectives are to ensure safe, efficient, and cost-effective flights through the use air and ground facilities [3]. Figure 1 shows the structure of ATM and explains the relations between ATM, ATS, and ATC.

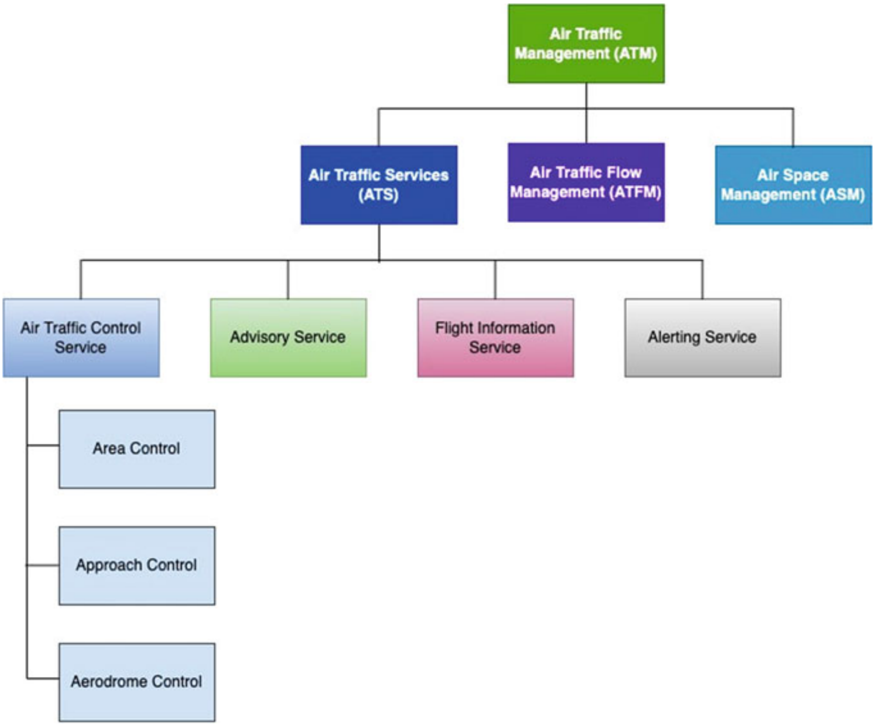


Fig. 1 ATM structure [4]

Airspace management (ASM) entails the planning, organization, and publication of air routes and control zones to ensure the safety of aircraft operations. Air traffic flow management (ATFM) contributes to regulating air traffic volume in accordance with airport and route capacity. Air traffic services (ATSs) are real-time services that separate air traffic to ensure safe takeoff, flight, and landing operations [4].

Flight information services, alerting services, and air traffic control are included in air traffic services (ATS). Flight information services provide essential data and recommendations for safe operating of aircraft. Alerting services alert the related agencies when an aircraft is in distress, need assistance, and support these agencies throughout the process. Air traffic control avoids incidents between aircraft as well as aircraft and maneuvering area impediments to expedite and maintain a controlled movement of aircraft [4].

Air traffic control (ATC) services are provided by three types of air control centers based on the phases of a flight. These phases of a flight are indeed the movements of an aircraft on the maneuvering area of an airport, including taxiing, landing, and takeoff, as well as the enroute cruising between arrival and departure. Air traffic controllers perform air traffic control services to a single area control-controlled flight via three distinct control facilities. An ATC tower at an airport is responsible for all air traffics inside the airport's maneuvering area [5]. Approach control centers offer air traffic control services to incoming and departing planes at an airport, and finally, the area control centers provide services to aircraft sailing in control zones. Throughout each phase of a controlled flight, these control centers offer well-ordered and methodical air traffic services.

For operations in ATC tower, the air traffic controllers perform their tasks by having visual contact with an aircraft in the maneuvering area and depend on the ATM system which provide information that has been fused from surveillance radars, alerting systems as well as related information to support their operations [5]. As for the other type of control facilities, the air traffic controllers will depend solely on ATM system alone without having visual contact with the aircraft.

2.2 ATM System

The air traffic management system (ATM system) is a system infrastructure which consists of multiple ICT hardware and software components that ingest, fused and process data from multiple sources such as surveillance sensors and messaging system and used it to provide information and management services to the air traffic controller to perform air traffic management [6]. It is also known as a platform because of its capacity to facilitate the coordinated integration of people, data, tools, and infrastructure with the help of communications, navigation, and surveillance systems stationed in the air, on the ground, or in outer space [6]. Common elements in an ATM system includes the following:

- Conflict management.
- ATM system delivery management.
- Traffic synchronization.
- Demand capacity balancing.
- Aerodrome operations.
- Airspace organizations and management.
- Airspace user operations.

These elements must be able to support all stages of air traffic management action which include strategic, pre-tactical, and tactical. Figure 2 shows the relation between elements and ATM action stages.

Strategic level action involves long-term information such as routes, flight slots, navigational, and communication facilities. Pre-tactical level action involves mid to near-term information requirement such as flight approval, notice to airmen (NOTAM), and meteorology forecast (MET), and finally, tactical level action involves all information requirement for real-time ATM operations such as trajectory update, emergency declarations, and other real-time changes involving day-to-day air traffic control operations [7]. All these information will be ingested by the ATM system and will be fused to form knowledge that is needed by the air traffic controllers.

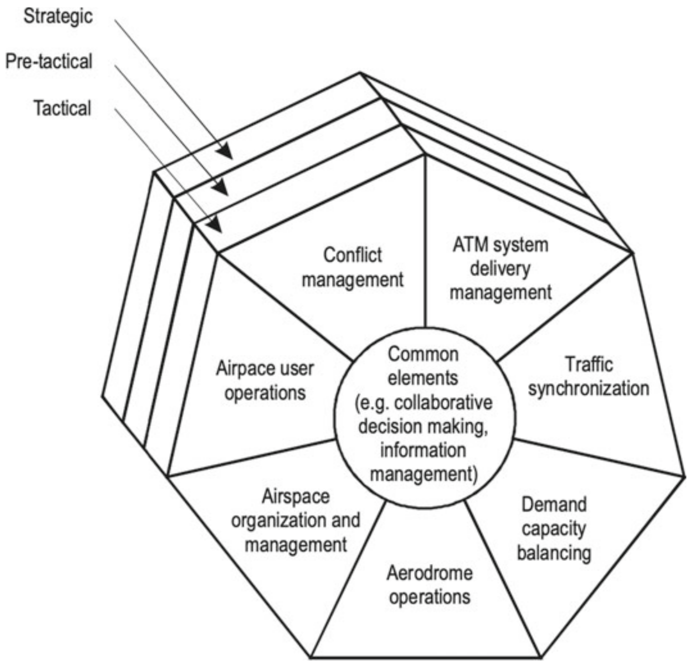


Fig. 2 ATM system elements [4]

2.3 Information Type in ATM

The lifeblood of air traffic management is information and it resulted from data sourced from multiple sub-system which is processed to form usable information for the operation of air traffic management [8]. The information will then become knowledge to the air traffic controllers and is used to make informed decision with regard to safety and efficiency of the air traffic services [9]. Type of information that is exchanged in ATM can be grouped as follows:

- *Aeronautical Information*—published as an Aeronautical Information Publication (AIP) by the local Aeronautical Information Service (AIS). Comprises information about the network, such as the capabilities that are available in the area, aeronautical charts (for example, airports), and navigational aids, among other things.
- *Flight Information*—air traffic service messages, which are exchanged utilizing an Aeronautical Fixed Telecommunication Network (AFTN). Contains all instantaneous updates to air crew, flight plan, location information, and flight changes.
- *Airport Information*—information about the airport's configurations, airside facilities, terminal maneuvering area (TMA), hazards and obstructions, approach profile, and other navigation and communication-related data. Also known as aerodrome information.

- *Weather Information (METAR)*—includes information such as significant meteorological events (SIGMET), terminal forecasts (TAF), and other relevant warning and awareness information that is deemed critical to flight operations. Originate from the national meteorology agencies of the country.
- *Flow and Capacity Demand Information*—generated by the network manager of flight operators and ANSPs using data from within and neighboring FIRs. These data are utilized to balance the flight operation network and manage traffic flow for the controllers.
- *Surveillance Information*—originated from radar, ADS-B, MLAT sensors, as well as other surveillance systems such as satellite-based surveillance. Plot or track of the traffic, which includes its speed, flight level, transponder information, and other associated codes, the information also includes additional types of codes.

Except for the surveillance information, this information is mostly available only in text form which uses the International Alphabet #5 (IA-5) or International Telegraphic Alphabet #2 (ITA-2) for transmission using the Aeronautical Fixed Telecommunication Network (AFTN) system as main platform of information sharing. Text forms are good for manual information processing; however, for machine-to-machine processing, a more modern and efficient format such as XML is required [10].

3 Identified Business Process in ATM

3.1 Background

With the increase of air traffic volume, data and information being produced by systems and sub-systems within the ATM ecosystem have also increased [11]. This situation requires systematic way to manage the flow and usability of the information. Machine-to-machine processing which enables automation is the way forward that needs to be undertaken to address this issue. ICAO under Aviation System Block Upgrade (ASBU) program has highlighted this issue and provided the guideline under the System Wide Information Management (SWIM) initiatives which can be adopted by Air Navigation Service Provider (ANSP) such as CAAM as the basis to address the issue [11].

3.2 Analysis Methodology

To understand the business process in ATM within CAAM, literature review was conducted on documents and manual involved in the day-to-day operation of ATM in CAAM. Documents and manual that have been examined are as follows:

- Aeronautical Information Publications (AIP) Malaysia (AIP AMDT 03/2022).

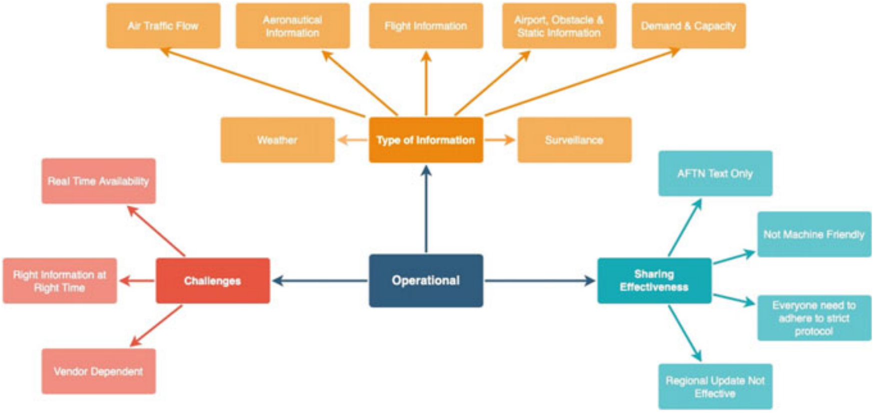


Fig. 3 Operational SME brainstorming mind map

- ICAO Annex 15—Aeronautical Information Services.
- ICAO Annex 4—Aeronautical Charts.
- ICAO Doc 8126—AIS Manual.
- ICAO Doc 8697—Aeronautical Chart Manual.
- ICAO Doc 10066—Procedures for Air Navigation Services—Aeronautical Information Management (PANS-AIM).

On top of the literature review, brainstorming methodology was also used. The brainstorming session was participated by subject matter experts (SME) from the air navigation service provider (ANSP) as well as the system provider which supplied the current ATM systems. Also involve are SMEs from the infrastructure provider which provide the system integration services. The findings from the brainstorming session are documented by using mind map methodology as per Fig. 3.

The following sections will explain the identified business process based on findings from literature review and the brainstorming session. This paper focusses on the top three business processes, namely Flight Planning, METAR, and NOTAM processes, which is heavily related to the day-to-day operation of ATM. All identified business processes have been documented by using the Business Process Model and Notation™ (Version 2.0) (BPMN™ 2.0) and presented again to the subject matter expert to get their endorsement in terms of the established as–if processes.

3.3 Business Process #1—Flight Planning and Distribution

A flight plan (FPL) is a document that provides air traffic service units with specific information on an aircraft's planned trip or flight segment. The ICAO Annex 2—Rules of the Air [12] and national flight information publications offer detailed rules surrounding the submission, contents, completion, modifications, and closure of a

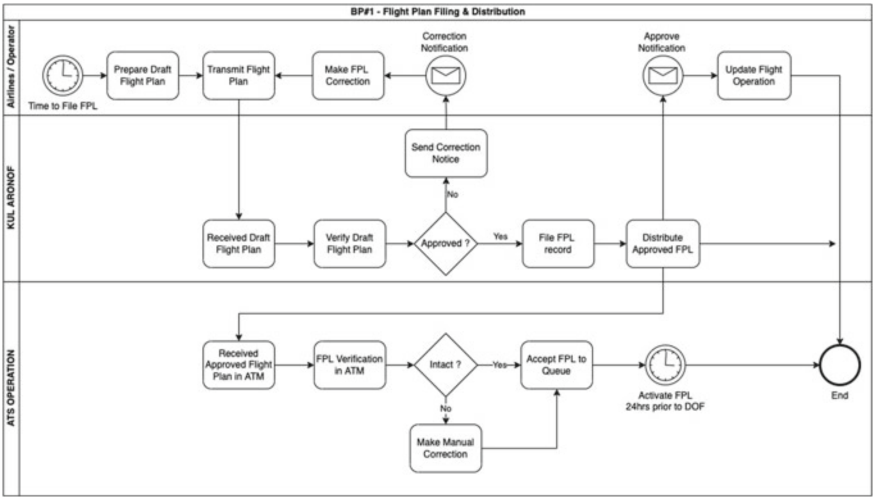


Fig. 4 Flight plan filing and distribution business process flow

flight plan. A flight plan may be submitted in the form of a written document, an electronic document, or orally. If a flight plan is required, it must be submitted prior to departure to an air traffic services reporting office or transmitted during flight to the appropriate air traffic services unit or air-ground control radio station, unless arrangements have been made for the submission of repetitive flight plans (RPLs).

As indicated in the AIP, in filing the flight plan, creator of the draft is airline, or the operator and the completed draft are submitted to the Kuala Lumpur (KUL) ARONOF for verification and approval. Flight plan that has been approved will be distributed to ATS operation unit as well as to the originator. The approved flight plan then will be processed by ATM system to get it ready for Air Traffic Services (ATS) operation. The flight plan has usually filed more than 24 h prior to the flight. The approved flight plan will be activated within 24 h prior to the date of flight. Figure 4 shows the business process flow of flight plan filing and distribution.

3.4 Business Process #2—MET Message Distribution

The meteorology services in Malaysia are provided by the Meteorological Department of Malaysia. For aviation specific services, it is provided by Aerodrome Meteorological Office (AMO) KLIA as Meteorological Watch Office (MWO) and a National Aviation Meteorological Centre. The meteorological services are provided for Air Traffic Services (ATS) Unit in the Kuala Lumpur FIR, as well as other AMO and Aeronautical Meteorological Station (AMS) in Peninsula Malaysia, including AMO Kota Kinabalu and AMO Kuching. The source of meteorology information originates from various sensor and forecasting tools such as Automatic Weather

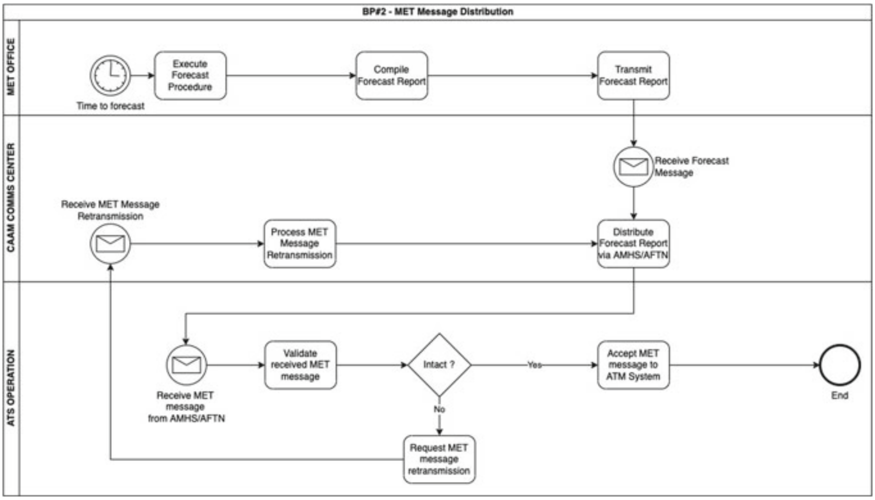


Fig. 5 MET message distribution business process flow

System at Aeronautical Meteorological Station, Automatic Weather Observing System (AWOS) along with Runway Visual Range (RVR) measurement adjacent to touchdown zones, mid-point, and stop-end for all runways, Terminal Doppler Radar (TDR) monitor of severe weather and wind shear, and upper-air observations at aeronautical meteorological station four times daily. This information is collected by the MET information fusion server and distributed to all users by using AFTN/AMHS platform. Figure 5 shows the process flow of MET messages for ATM system usage.

3.5 Business Process #3—NOTAM Management

Understanding the installation, condition, or change in any aeronautical facility, service, procedure, or danger as soon as possible is a top priority for all employees involved in flight operations, and this is exactly what NOTAMs provide. Each NOTAM follows the format specified by the ICAO NOTAM Code and includes ICAO acronyms, indications, identifiers, designators, call signs, frequencies, figures, and plain English. The NOTAM Office publishes and disseminates NOTAM for the Kuala Lumpur and Kota Kinabalu FIRs in the following four series:

- Series A—KUALA LUMPUR FIR for International distribution.
- Series C—KUALA LUMPUR FIR for Domestic distribution.
- Series D—KOTA KINABALU FIR for International distribution.

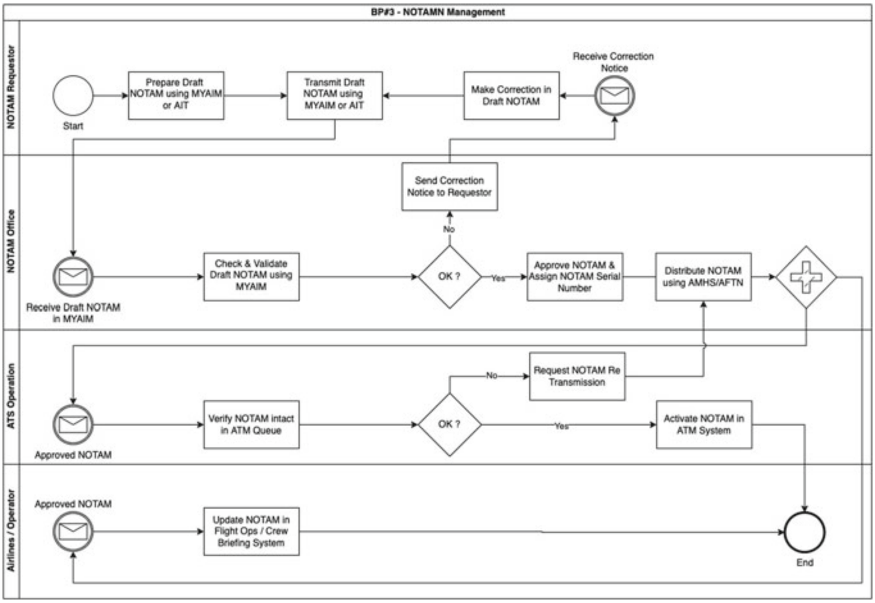


Fig. 6 NOTAM management business process flow

- Series F—KOTA KINABALU FIR for Domestic distribution.

With the use of MYAIM, aeronautical data for pre-flight briefing services may be accessed and managed with ease. Pilot Briefing Offices at KLIA, Sepang, Sultan Abdul Aziz Shah Airport, Subang Control Tower, Langkawi International Airport, Penang International Airport, Senai International Airport, Kota Kinabalu International Airport, Kuching International Airport, and Miri Airport also have MYAIM Terminals. Printouts of the briefing, complete with all pertinent NOTAM data, are accessible on demand through printers connected to the MYAIM terminals, which may be accessed either on-screen or remotely. Interfacing the MYAIM system with AFTN/AMHS allows NOTAM messages to be sent to stations without MYAIM terminals, as well as to other ANSPs across the globe and the ATM system. Figure 6 shows the NOTAM Management business process flow.

4 Discussion and Way Forward

The understanding and documentation of current business processes in ATM information exchange is important to ensure that all information components are captured during the development of the information exchange interoperability framework. During the research, other business process which comprises support functions process flow will also be documented and studied. A well-rounded and comprehensive understanding of all information components and its related business process is one of the success factors to a practical and effective framework. It will also help to understand the main platform and technology that is currently use for ATM information exchange.

From the analysis of facts gathered in this study, it can be understood that the current platform for information exchange for ATM system is the Aeronautical Fixed Telecommunication Network (AFTN) and the ATS Message Handling System (AMHS). Except for the surveillance sub-system, all other information sources use this platform to exchange information with the ATM system. Majority of the information is being exchanged by using the IA-5 character set with a specific message formatting in accordance with each information services such as FPL, METAR, or NOTAM. The information that contains in these messages are human readable and made for manual human processing. However, the capability in processing these messages efficiently is limited by the nature of serial character processing and the processing and translation of these messages heavily dependent on the parser design of each sub-system that produce or consume the information.

Implementation of initiative to address the limitation of serial data processing by using a predefined XML messaging format as defined in the ICAO SWIM guideline will address the processing performance issue. At the same time, the predefined XML messaging format will also address the interoperability issue between systems by ensuring transparency in information sharing protocol. A "vendor-lock" situation can be greatly reduced in implementing new information services related to ATM.

However, to implement SWIM in the Malaysia's ATM environment, a framework that addresses information exchange interoperability is required to ensure that the implementation of SWIM infrastructure, which will be unique to Malaysia is correctly designed, developed, and operated by all stakeholders. With the understanding of core business processes in the ATM information exchange, the framework that will be developed will be accurate, logical, and practical to use by the stakeholders.

References

1. Graham W, Hall C, Morales M (2014) The potential of future aircraft technology for noise and pollutant emissions reduction. Transport Policy 36–51
2. IATA (2019) Aviation cyber security rountable. IATA, Singapore
3. SKYbrary (2017) SKYbrary. Retrieved from air traffic management definition: https://skybrary.aero/articles/air-traffic-management-atm
4. ICAO (2005) Doc 9854-AN/458—Global ATM operational concept. ICAO, Montreal
5. Arblaster M (2018) Air traffic management economics, regulation and governance. Elsevier
6. ICAO (2016) Doc 9750-AN/963—Global Air Navigation Plan 2016–2030. ICAO, Montreal
7. ICAO (2016) Procedures for air navigation services (PANS)—air traffic management (Doc 4444). ICAO, Montreal, Canada
8. Lootens KJ, Efthymiou M (2019) The adoption of network-centric data sharing in air traffic management. Inf Resour Manag J 32(3):48–69

9. Mondoloni S, Rozen N (2020) Aircraft trajectory prediction and synchronization for air traffic management applications. Progress Aerosp Sci 119
10. Bellamy III W (2014) Avionics Big Data: Impacting All Segments of the Aviation Industry. Retrieved from aviation today: https://www.aviationtoday.com/2014/12/01/avionics-big-data-impacting-all-segments-of-the-aviation-industry/
11. ICAO (2018) Manual on system wide information management (SWIM) concept. ICAO, Montreal
12. ICAO (2005) Annex 2: rules of the air, vol 10th edn. ICAO, Montreal, Canada

New Method for Generating a Regular Polygon

Penio Dimitrov Lebamovski

Abstract This paper presents a new method for generating a regular polygon. It is based on the method of limits of Isaac Newton and the method of indivisibles of the Italian mathematician Boenoventura Cavalieri. The traditional way to construct a regular polygon is based on trigonometry, which uses (sin, cos, and radius), respectively. Several vertices and radius characterize it. The new method defines the polygon by the number of vertices and the side length. And it only uses relationships of parallel segments. The new way allows drawing a polygon with a number of vertices stretching to infinity. This is one of the main differences between the two methods, which is significant. In the existing method until now, the number of vertices is limited to a specific value. Through the new way, the number of vertices can grow to infinity. It can be a disadvantage because it requires a lot of computer power. These new polygons can form complex geometric shapes, such as polyhedra (prism, truncated pyramid, and pyramid). Thanks to the new method, the polyhedra are mathematically more accurate than the traditional way, which should extrude the polygons to polyhedra.

Keywords Boundary method · Polyhedron · Regular polygon · 3D software · 3D technology

1 Introduction

In the theory of mathematical education, which deals with spatial and geometric imagination development, several experts and researchers are united in their opinion that these imaginations are poorly developed in some students [1]. For the solution to this serious problem faced by modern education, especially in schools and universities, 3D technology comes to the aid of educators, which can, in turn, include: systems for virtual environments, such as stereo visualization systems with and

P. D. Lebamovski (✉)
Institute of Robotics, Bulgarian Academy of Sciences, Sofia, Bulgaria
e-mail: p.lebamovski@ir.bas.bg

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_75

without immersion. An example of immersive systems which can be used in school is a virtual reality helmet and the high-budget Wedge system. The mentioned technologies should enable students to study the geometrical objects along the three dimensions, which, as is known, value along (abscissa, ordinate, and z—coordinate). At the same time, they should allow for all kinds of manipulations. A large part of the geometric objects that are studied in geometry is located in 3D space. Where it is necessary to develop special qualities in students, such as: imagination in 3D geometry, logical thinking, and practical understanding of the taught material. Some of the most famous 3D software systems which help teachers to achieve excellent learning results are GEOGEBRA, CABRI 3D, DALEST, etc. These systems open up new opportunities for the educational process. Of the mentioned software, only one of them allows the use of 3D visualization technology and that is GEOGEBRA, which uses passive anaglyph projection for visualization. However, if there is a suitable method for constructing geometric objects, in particular by stereometry. It will be possible to develop spatial imagination in students who still need it. If the learning process is turned into a game, then great results would be achieved in student learning. If more virtual, augmented and mixed reality devices would be used, including 3D printer technology. Then, the result of teaching geometry would be better. Therefore, this paper presents a new method by which immersive and non-immersive virtual reality systems can visualize stereometric geometric figures.

The purpose of this paper is to present a new method for generating a regular polygon. Based on it, more complex geometric objects than stereometry can be constructed. This new method gives a much better result than the traditional method.

2 Methodology

2.1 Cavalieri's Method

Before the exposition of the method, Cavalieri says the following: "Whether the continuous consists of indivisibles or not, the sets of indivisibles are comparable to each other, and their magnitudes are in a certain relation to each other" [2]. Cavalieri's method is to compare figures' faces through all possible lines, which we present as sections of the figures with lines that move and are parallel to the rule (direction) all the time. By analogy, the totality of all their parallel sections is examined when comparing the volumes of bodies. If we have two figures of equal size, all their possible lines are equal. Regardless of the rule, we have chosen. It follows that all the indivisibles of a given figure taken under some arbitrary rule are equal to all the indivisibles of the same figure taken under any other rule. Figures are related to each other as an arbitrary rule, and solids take all their lines as all their planes taken by an arbitrary rule. The statement that Cavalieri considers fundamental is that to discover the relations between two planes or spatial figures, it is sufficient to find the relation of all indivisibles according to a given rule. The statement that is included in geometry textbooks reads as follows: The faces or volumes of two figures are equal if and only if the faces or the lengths of all their corresponding sections parallel to a given plane or line are equal to each other. 26 In short, the “method of indivisibles” is used to determine volumes and faces of surfaces by many parallel lines and planes, also known as Cavalieri’s Principle.

2.2 Newton’s Method

In a part of his book, teaching immediately after the lemmas, Newton compares the method of limits with the “method of indivisibles” of Cavalieri. Newton shared the following thought—“I set forth the preceding lemmas to avoid long boring proofs with a contradiction in the manner of the ancients”. The proofs by the method of indivisibles are also shortened, but this method is less geometric and creates more significant difficulties in using you. In truth, the method of limits gives the same results as the “method of indivisibles”. The following statement arises from the fact that the last relations of the vanishing quantities exist, and the vast quantities of the indivisibles also exist. To this opinion, Newton replies as follows: the last relations by which the magnitudes vanish are not relations of the last magnitudes, but limits to which the relations of diminishing magnitudes are all the time approaching, which may be approached more than a given difference, provided that they do not can neither surpass nor reach them before the magnitudes diminish infinitely. Different differential (infinitesimal) representations were so effective that Newton never gave them up. However, over time, the use of infinitesimal quantities—be they indivisibles, quantities smaller than an arbitrarily given quantity, or extremely small—is not rigorous enough. In connection with the abandonment of infinitesimally small addendums, Newton wrote: “In mathematical matters, the smallest errors must not be overlooked”. Newton was the first to introduce the term “boundary”, but he did not give any definition of the concept of “boundary” and its properties, which reduces the importance of the proofs of the boundary transition theorems. Newton also considered this concept intuitive [2, 3]. The method of limits was set forth by Newton in twelve lemmas. In this article, to clarify the concept of limit, the following two lemmas of Newton are considered:

1. Lemma 1: The graphical representation of the lemma is shown in Fig. 1 [3]. An arbitrary geometric figure AacE is chosen, which is bounded by the straight lines Aa and AE and the curve acE, and they fit into any number of rectangles with diagonals: Ab, Bc, Cd, Do, which have equal bases: AB, BC, CD and sides: Bb, Cc, Dd. Reducing the length of the sides of the rectangles: aKbl, bLcm, cMdn etc. and increasing their number to infinity, then, according to Newton, the limit of the relations of the figure AKbLcMdD, the circumscribed figure AalbmcndoE and the curvilinear figure AabcdE are in the ratio 1:1:1 [3].

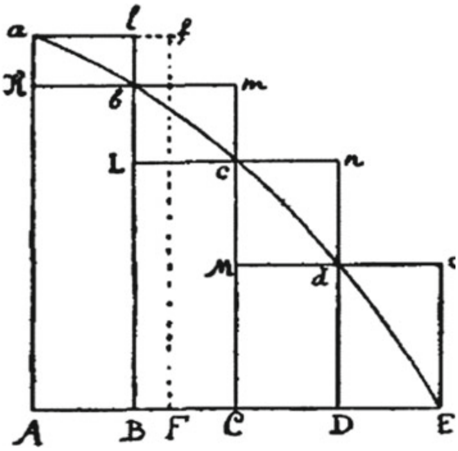


Fig. 1 Lemma newton [3]

2. Lemma 2: For similar figures, the lengths of the corresponding sides, both rectilinear and curvilinear, are proportional, and the faces of the figures are proportional to the squares of the sides [3].

In the eighteenth century, his theory of limits found critics who considered it logically imprecise and commentators who disputed the meaning of its applications, but some advocates developed the theory, such as Newton. Only in the 1920s, Cauchy began a complete synthesis of Newton's ideas, which is still the basis of mathematical analysis. Newton's lemmas listed are of immense importance. It does not give a uniqueness theorem on a limit or between a limit and an infinitesimal variable.

3 Results and New Method

The traditional method of generating a regular polygon is based on trigonometry. Several vertices and radius characterized it. To calculate the length of its side, the relationship between the length of the side and the radius of the inscribed or circumscribed polygon is used. In order to be able to generate a prism or a pyramid, it is necessary to use a technique from three-dimensional computer graphics known as extrusion. This process is, unfortunately, not very suitable for stereometry applications. For this purpose, this article presents a new way with much greater accuracy in drawing polywalls than the mentioned technique. The new method is characterized by the number of vertices and the length of the side of the regular polygon with side **a** [4, 5]. And here, as an additional parameter, the radius of a circle inscribed or circumscribed around the polygon can be calculated. Using the relationship between the side of the polygon and the radius. This innovative method is mathematically more accurate than the traditional one, as it is unnecessary to extrude a 2D polygon based on trigonometry. It only uses relationships of parallel segments, not trigonometry.

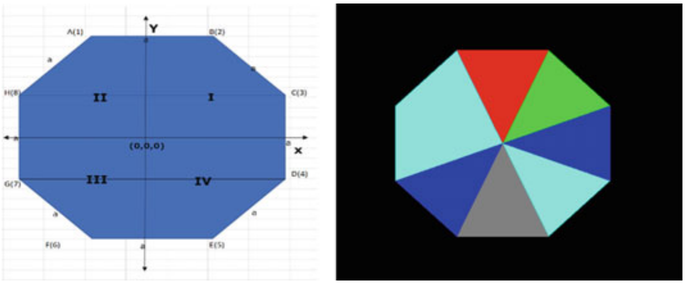


Fig. 2 Regular polygon

The disadvantage is that it requires a lot of computing power. Knowing the length of the side of a regular polygon, the values along the abscissa and ordinate of its vertices can be calculated.

During the calculation itself, parallel segments and their relations along the two axes are used. The regular polygon is placed at the center of the 3D coordinate axis Fig. 2, with z coordinate values equal to zero for each of the vertices of the base. Next is the determination of the relations of the parallel segments along the abscissa and ordinate; in this case, they are 1:2:3:4 = 1:2.34:2.34:1. The polygon is divided into three parts: an isosceles trapezoid, a rectangle, and another isosceles trapezoid. The height of the first figure is involved in calculating the first and second vertices, respectively, along the ordinate; the value for the abscissa is equal to **a/2**. The value for the ordinate at the third and fourth vertices is equal to **a/2,** and that for the abscissa is half of the lower base of the isosceles trapezoid. The calculation of the values of the fifth and sixth peaks is similar to those of the first and second peaks. The difference is along the ordinate, where in this case, the values will be negative. The calculation of the seventh and eighth vertices is similar to these values for the third and fourth vertices. Here, there will be negative values on the x-axis. Complex geometric figures such as polygons, pyramids, and prisms can be constructed based on this regular polygon.

To draw a pyramid Fig. 4 and Table 2, one more vertex showing the pyramid's height is needed. Here, there are two more parameters **h1** and **h2**, if they have a value equal to zero then the final result will be a straight pyramid. If one of the values is zero, and the other is a non-zero number, the end result will be a tilted pyramid. A regular octagonal prism can be constructed analogously Fig. 3 and Table 1. Here, it is necessary to add another regular polygon, as the upper base of the prism. And here, there are also added two additional parameters **h1** and **h2**, defining a straight and an inclined prism. The polygons created by the new boundary method are much more effective than those made by trigonometry (Table 2).

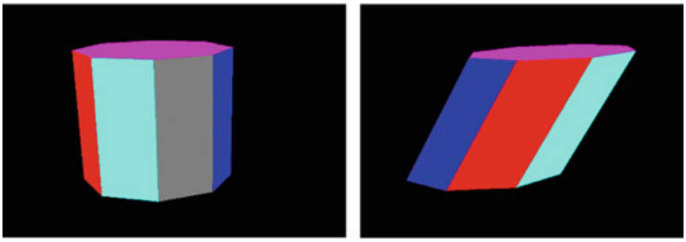
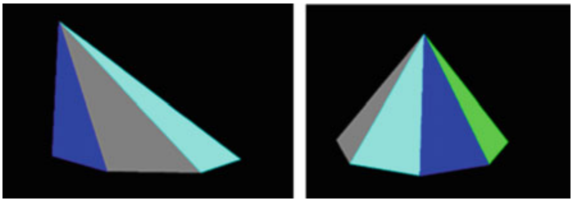


Fig. 3 Prism with 8 vertices

Fig. 4 Pyramid



They have many advantages, such as

1. Visualization through virtual augmented and mixed realities
2. They are more flexible and can be easily manipulated, for example, coloring the walls of a polyhedron

The traditional way to construct polyhedra is to extrude a polygon characterized by the number of vertices and radius. Extrusion is a technique used in 3D graphics. The principle of this method is that by moving a 2D graphic, complex three-dimensional shapes can be formed. Two-dimensional graphics transform by translation, rotation, or movement along a arbitrary curve. This technique is used by geometry software, not only for drawing 3D models and stereometric figures. The new method uses only number relations and is based on Cavalieri's method of indivisibles and Isaac Newton's method of limits. For example, the volume of a cylinder, cone, and sphere can be calculated using the method of limits. For example, the volume of a cylinder is called the limit to which the series of the volumes of the regular prisms inscribed in it tends. In this method, the number of prism walls grows to infinity. Using the new boundary method, a regular polygon can be drawn with the number of its vertices growing to infinity.

Similar to Newton's method of limits, through the new approach, to reach the boundary of the studied figure (a regular polygon), the values in the case of the number of vertices must be increased to infinity. The goal is to reach the formation of a circle. The essence of Cavalieri's method is that the geometric figure can be divided by the sections used parallel to a given rule, which can be segments or planes in the two-dimensional and three-dimensional cases, respectively. Like the method of Cavalieri's Indivisibles with the new process, the geometric figure can be divided by passing parallel sections. And from there, it follows that these sections are in a

Table 1 Prism with side of base equal to **a**

Number of vertex	X	Y	Z
1	$-a/2$	$a * \mathrm{sqrt}(0.5511) + a/2$	0
2	$a/2$	$a * \mathrm{sqrt}(0.5511) + a/2$	0
3	$a/2 + 0.67 * a$	$a/2$	0
4	$a/2 + 0.67 * a$	$-a/2$	0
5	$a/2$	$-(a * \mathrm{sqrt}(0.5511) + a/2)$	0
6	$-a/2$	$-(a * \mathrm{sqrt}(0.5511) + a/2)$	0
7	$-(a/2 + 0.67 * a)$	$-a/2$	0
8	$-(a/2 + 0.67 * a)$	$a/2$	0
9	$-a/2 - h1$	$a * \mathrm{sqrt}(0.5511) + a/2 - h2$	h
10	$a/2 - h1$	$a * \mathrm{sqrt}(0.5511) + a/2 - h2$	h
11	$a/2 + 0.67 * a - h1$	$a/2 - h2$	h
12	$a/2 + 0.67 * a - h1$	$-a/2 - h2$	h
13	$\mathrm{a}/2 - h1$	$-(a * \mathrm{sqrt}(0.5511) + a/2)- h2$	h
14	$-\mathrm{a}/2 - h1$	$-(a * \mathrm{sqrt}(0.5511) + a/2)-- h2$	h
15	$-(a/2 + 0.67 * a)\text{-}h1$	$-a/2 - h2$	h
16	$-(a/2 + 0.67 * a) - h1$	$a/2 - h2$	h

specific relation to each other. In the case of a regular octagon, the boundary relations of four parallel segments are used. In this case, they are 1:2:3:4 = 1:2.34:2.34:1. The main contribution of the new approach is that it gives a more accurate result than the traditional one using trigonometry. When constructing a polyhedron, it is the best alternative. The conventional way to build a polywall is by extruding polygons. The disadvantage of the new approach is that it requires too much computing power. In rare cases, an error of about 0.05% can be reached in determining the ratios of parallel segments.

Table 2 Pyramid with side of base equal to **a**

Number of vertex	x	y	z	Wall
1	$-a/2$	$a * \text{sqrt}(0.5511) + a/2$	0	1,2,3,4,5,6,7,8
2	$a/2$	$a * \text{sqrt}(0.5511) + a/2$	0	1,2,9
3	$a/2 + 0.67 * \text{a}$	$a/2$	0	2,3,9
4	$a/2 + 0.67 * \text{a}$	$-a/2$	0	3,4,9
5	$a/2$	$-(a * \text{sqrt}(0.5511) + a/2)$	0	4,5,9
6	$-a/2$	$-(a * \text{sqrt}(0.5511) + a/2)$	0	5,6,9
7	$-(a/2 + 0.67 * a)$	$-a/2$	0	6,7,9
8	$-(a/2 + 0.67 * a)$	$a/2$	0	7,8,9
9	$h1$	$h2$	h	8,1,9

4 Conclusion

This paper presents a new way to generate a regular polygon based on Cavalieri's method of indivisibles and Isaac Newton's method of limits. On its basis, complex geometric shapes can be created, such as polyhedron (pyramid and prism). It only uses a relationship between numbers, not trigonometry. 2D and 3D graphing software use a traditional method that is based on trigonometry. But they have a limit on the number of vertices, about some values. The new way allows vertices to grow without limit (to infinity). Here, it is only necessary to determine the relationship of sections along the abscissa and ordinate. This is necessary to calculate the values of the vertices along the three dimensions. Future work will focus on 3D game development and how they can participate in the analysis of cardiac data obtained from a Holter device. In the 3D modeling of various shapes, the polygons proposed in this article, innovative for programming, will be used.

References

1. Rahman MHA, Puteh M (2017) Learning trigonometry using GeoGebra learning module: are under achieve pupils motivated? AIP Conf Proc 1750(1):39–42. https://doi.org/10.1063/1.4954586
2. Bashmakova I (1975) Istoria na matematikata, tom № 2, Sofia, Nauka I Izkustvo
3. Newton I (2002) The mathematical principles of natural philosophy. In: Wilkins DR (ed)
4. Lebamovski P, Petkov E (2020) Usage of 3D technologies in stereometry training. CBU Int Conf Proc 1:139–146. https://doi.org/10.12955/pss.v1.61
5. Lebamovski P (2021) The effect of 3D technologies in stereometry training. CBU Int Conf Proc 1:68–74. https://doi.org/10.12955/pns.v2.155

Method for Eliciting Requirements in the Area of Digital Sovereignty (MERDigS)

Maria Weinreuter, Sascha Alpers, and Andreas Oberweis

Abstract Digital sovereignty has become increasingly important in socio-political discourse owing to the increased perception of the absence of the state of digital sovereignty. This state is based on the contradictory requirements of various parties caused by the heterogeneity of stakeholders, holding various needs and desires. A method of eliciting requirements in the area of digital sovereignty (MERDigS) was developed to create a noticeable requirements basis for software development projects that intend to enable their stakeholders in the state of digital sovereignty. It can be used in software development projects to elicit the requirements of stakeholder groups in isolation. MERDigS adopts a human-centred approach, enabling it to address the needs and desires of stakeholders as the source of requirements in digital sovereignty. It captures the full range of requirement types by implementing modules to elicit noncommunicable requirements. MERDigS is developed by adapting comparative work in the field of requirements engineering. Three experts were interviewed to assess the plausibility of the MERDigS approach. The assessment showed that this approach is plausible and reasonable. In addition, MERDigS can be designed to be more generic so that its implementation can be easily adapted to various software development projects. Future work might incorporate a parallel exchange with developers into MERDigS to directly discuss technical implementation options for elicited requirements. Further incorporation might moderate the requirements between different stakeholder groups.

Keywords Digital sovereignty · Requirements elicitation · Human-centred approach · E-Governance and government · Value-based requirements engineering

---

M. Weinreuter (✉) · S. Alpers · A. Oberweis
FZI Forschungszentrum Informatik, 76131 Karlsruhe, Germany
e-mail: weinreuter@fzi.de

S. Alpers
e-mail: alpers@fzi.de

A. Oberweis
e-mail: oberweis@fzi.de

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_76

1 Introduction

The attainment of a state of digital sovereignty is becoming increasingly important in socio-political discourse. Digital sovereignty refers to a state in which stakeholders are in full control and have freedom over their conscious actions and decisions in the digital space. Stakeholders seeking and achieving digital sovereignty range from individuals to organisations, civil societies, states, and confederations of states, providing it a wide range of meanings [1–3]. The state of digital sovereignty can be attained under four general criteria regardless of the diverse emphases of meaning.

1. *Knowledge of the operation of digital technologies.* This is a necessary basis for achieving digital sovereignty. With this knowledge, the potential consequences and implications of use should be understood and applied [4].
2. *Choice between different alternatives.* A choice should be made between various alternatives [5]. On the one hand, if no choice is made between different alternatives with comparable capabilities, digital technologies should be developed and produced by the stakeholder. On the other hand, they should be developed and produced by the stakeholder if an alternative represents a key technology for the stakeholder.
3. *Possibility to decide and act according to interests and competencies.* Stakeholders should develop their digital space according to their interests and competencies and individually shape their digital environment. [Cf. 6, 7]
4. *Control in dealing with digital technologies.* Control refers to the ability to influence the use of digital technologies obtained from external providers, determine their dynamics and impact, and check and correct deviations [Cf. 7, 8].

Although these criteria can be defined in general terms, stakeholder heterogeneity leads to different manifestations of these criteria. The stakeholder determines, for example, digital technologies that are designated as key technologies, external providers who can be trusted, or the extent to which the further processing of data must be transparent so that they can act and decide confidently. This lack of clarity motivates us to pursue our desire to attain the state of digital sovereignty. To this end, this study introduces a method for eliciting requirements in the area of digital sovereignty (MERDigS). Software engineering requirement elicitation methods are transferred and adapted to digital sovereignty requirement elicitation methods to develop MERDigS. The first and most important phase of requirements engineering is requirements elicitation. Requirements elicitation gathers information about the requirements and context of a project [9, 10]. The information is collected through appropriate elicitation methods, either directly through stakeholders or indirectly through other requirements sources [9, 11].

The research methodology that explains the development process of MERDigS is outlined below. Hence, MERDigS is described through the field of application, its artefacts, and the process model steps. The fourth chapter evaluates MERDigS using three expert interviews to discuss the plausibility of the process model in MERDigS. Finally, a brief conclusion and outlook for future work are presented. A detailed presentation of the method is published as a companion white paper [12].

2 Research Method

This study was conducted in two steps. First, a variety of requirements elicitation techniques, including their advantages and disadvantages, were considered in the form of a broad literature search. Simultaneously, highly relevant comparative works, whose approaches could be adopted by MERDigS to a large extent, were examined. Consequently, a potential set of elicitation techniques can be identified. The search for potential elicitation techniques focused on various complexity dimensions that MERDigS must overcome. The selection of complexity dimensions was inspired by Angelis et al. [13]. The following complexity dimensions were derived for MERDigS:

1. *Stakeholder heterogeneity.* Complexity through stakeholder heterogeneity results from the lack of knowledge regarding the stakeholders addressed by the execution of MERDigS. Therefore, MERDigS must apply to all currently conceivable stakeholders of digital sovereignty.
2. *Project heterogeneity.* The fact that the selection of requirements elicitation techniques is generally dependent on the project contributes to complexity through project heterogeneity [11, 14]. Therefore, MERDigS must abstract from individual software development projects and provide a general generic approach.
3. *Communication.* Complexity through communication results from the inability of stakeholders to communicate all their needs in an understandable and communicable manner [15, 16]. In addition, MERDigS elicits both communicable and noncommunicable requirements.
4. *Abstractness.* Complexity through abstractness results from various factors such as desires, needs, values, insecurities, and fears of the stakeholders, which strongly influence requirements in the area of digital sovereignty. Owing to such factors, MERDigS requires more in-depth investigations than, for instance, the elicitation of functional requirements for general-purpose software. In addition, MERDigS seeks requirements for an artefact, which is a human construct and whose scope is difficult to grasp.
5. *Accessibility.* Complexity through accessibility arises because some elicitation techniques require experienced applicants [17]. Nevertheless, MERDigS must also apply to inexperienced applicants.
6. *Resource provision.* Complexity through resource provision arises from the fact that MERDigS does not elicit requirements for an entire project. MERDigS elicits requirements for an area that the project should cover. Therefore, a limited willingness to invest effort must be assumed. Concerning other requirements elicitation methods, MERDigS should therefore be feasible in a time-efficient manner.

7. *Transparency.* Complexity through transparency results from the implementation of MERDigS, followed by additional phases in which requirements are processed. Transparent documentation of elicited requirements is necessary to ensure that these phases can follow smoothly.

Different platforms and databases were systematically searched for peer-reviewed scientific articles to select comparative work, which was then evaluated using complexity dimensions. After ruling out the existence of any method that could specifically elicit requirements in the area of digital sovereignty, we defined the following search strings:

a. *(requirements elicitation)* $\wedge$ *(multiple stakeholders* $\vee$ *method* $\vee$ *process* $\vee$ *mixed methods* $\vee$ *ubiquitous systems* $\vee$ *embedded systems* $\vee$ *large project* $\vee$ *social topic* $\vee$ *agile method* $\vee$ *goal oriented* $\vee$ *collaborative method* $\vee$ *nonfunctional)*
b. *(digital sovereignty)* $\wedge$ *(method* $\vee$ *requirements)*

By querying these search strings, a total of 110 comparative papers were chosen on the first screen. After applying inclusion and exclusion criteria on a second screen, this number was reduced to 32. In particular, papers presenting methods for eliciting requirements in large-scale research projects and ubiquitous systems in the context of software development projects were included. Furthermore, papers focusing on stakeholders and their needs and desires, for instance by pursuing a value-based requirements elicitation approach or highlighting the direct communication with the stakeholders, were included [18, 19]. In addition, reviews of existing requirements elicitation methods were included. On the contrary, especially papers that only propose single tools were excluded. The reason for this at this point of time was the uncertainty if the individual method for which the tool is applicable is suitable at all. In addition, papers whose focus was not on requirements elicitation but on the entire requirements engineering process were excluded. Such papers provide, for example, methods for documenting and managing requirements or for negotiating conflicts between stakeholder groups. Following the selection of potential techniques, these were checked and improved using a predefined selection table. For this purpose, the selection table of Gupta and Deraman [20] was selected and applied to the parameters of MERDigS. Following these steps, four techniques were chosen: document analysis, focus groups, interviews, and questionnaires, which when combined could overcome the complex dimensions of MERDigS. Thus, in step two, more specific literature research on these techniques could be conducted, filling in missing modules and extending modules for the development of MERDigS.

3 Method for Eliciting Requirements in the Area of Digital Sovereignty (MERDigS)

MERDigS is used to elicit communicable and noncommunicable digital sovereignty requirements for a software development project. Moreover, it is used in the initial project requirements elicitation, wherein additional requirements arising from the area of digital sovereignty are collected. To this end, MERDigS adopts a human-centred approach, addressing the needs and desires of stakeholders [21]. The appropriate combination of elicitation techniques leads to an approach in which the relevant needs and desires of stakeholders can be elicited at deeper levels. The combination can be obtained from the steps shown in Fig. 1. Therefore, the stakeholders are considered at a (predominantly) abstract level at the beginning of MERDigS, so that basic requirements that apply to all stakeholders in the stakeholder group have already been collected and validated following the conduction of focus group-like workshops. By considering the basic requirements as completed, the subsequent conduction of the interviews can force the elicitation of stakeholder-specific quality requirements. Under this condition, the human-centred approach can be realised in the interviews, in which in-depth conversations about the needs and desires of the stakeholders are held.

3.1 Field of Application

MERDigS can be used for a wide range of software development projects. Applicants are representatives of software development projects in research institutions, companies, and social institutions, such as project managers, requirements engineering managers, or social technology designers. The application of MERDigS isolates the requirements of one stakeholder group. The stakeholder groups to which MERDigS can be applied are as follows: (1) natural persons, (2) a state, (3) public administration, (4) scientific organisations, (5) business organisations, and (6) civil

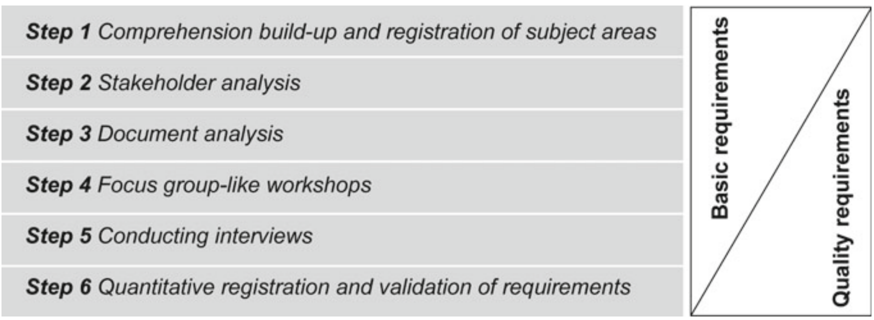


Fig. 1 Steps in MERDigS

society organisations. They represent the intersection of the stakeholders of digital sovereignty and typical stakeholders of software development projects.

3.2 Artefacts

During the execution of MERDigS, artefacts are generated, modified, and refined. The information of the intermediate results resulting from the execution of the individual steps is stored with these artefacts. The following steps can enable accessing these artefacts any time, generating new data for each access. Figure 2 shows the rough course of the information content of the most important artefacts in the MERDigS process model. Furthermore, it indicates artefacts that can/need not be prioritised in a given step, as well as the content interpretation and weighting at each step. Notably, users can choose design templates and application programmes when creating artefacts. However, the definitions of terms and word usage should adhere to the definitions of a glossary, which should be completed by step 3. (Sect. 3.3).

The most important artefacts in MERDigS are as follows:

1. *Requirements document.* The requirements document contains all relevant information used, collected, and generated during the execution of MERDigS. It is generated by MERDigS and serves as an input for subsequent requirements engineering phases. The final requirements register, into which project descriptions, stakeholder descriptions, nonextended topic areas, and other artefacts can be inserted, is the primary content of the requirements document [19].
2. *Requirements register.* A requirements register is a document in which the collected, not necessarily final, requirements are documented consistently and understandably [22, 23]. The existing requirements register for the software development project should be used as a template.

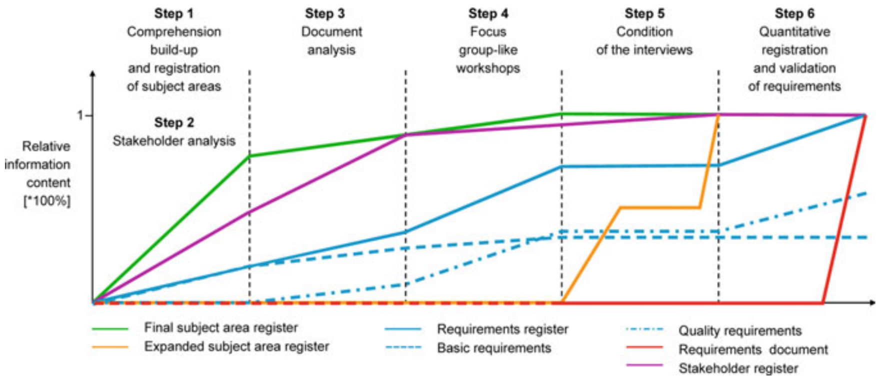


Fig. 2 Relative information content of the most important artefacts

3. *Stakeholder register*: Stakeholder types are stored in a stakeholder register with names, properties, needs, and desires. Furthermore, the proportion of stakeholders in a stakeholder list assigned to the stakeholder type can be added. If available, suitable stakeholder representatives of the stakeholder types are stored.
4. *Subject area register*: The subject area register stores subject areas in which the requirements in the area of digital sovereignty for a specific project can be collected and stored. Subject areas, subordinate subject areas, and requirement types are important contents of the subject area register. During the execution of MERDigS, the subject area register is expanded to the requirements register and ambiguous statements, subsequently known as the *expanded subject area register*.

3.3 Steps

The process model of MERDigS is divided into six steps that run sequentially. The six steps are as follows:

Step 1 Comprehension build-up and registration of subject areas First, digital sovereignty should be understood. Even if this understanding already exists, this activity is required to perceive the meaning of the emphasis on digital sovereignty that results from the viewpoint of a specific stakeholder group [24]. Following this, a subject area register, which serves as a basis for eliciting complete requirements, is established [23, 24, 26]. In particular, the subject area register helps collect all relevant needs and desires of stakeholders to subsequently derive them into requirements in the area of digital sovereignty [26, 27]. Table 1 can be used for identifying some, but not necessarily all, potential subject areas. Then, the subject area register is compared to the requirements already elicited for the project to ensure that no requirements are elicited twice. This makes room for MERDigS's human-centred, qualitative approach, which can be more responsive to stakeholders' needs and desires [26].

Step 2 Stakeholder analysis Stakeholder analysis involves developing an understanding of the stakeholders in a stakeholder group and storing it in a stakeholder register. Knowledge in building enables analysts to form a stronger bond with stakeholder representatives in subsequent steps, enabling them to better address their needs and desire [23, 30]. To this end, the stakeholders of the stakeholder group are first analysed individually, concerning the software development project, and then jointly [31]. Stakeholder types must group stakeholders with similar characteristics and must be diverse to the extent possible among themselves. Following this, information on stakeholder types is stored in a stakeholder register. Furthermore, stakeholder representatives from each stakeholder type are selected and recorded in the stakeholder register. Because each stakeholder type has similar stakeholder characteristics, it follows that they have similar needs and desires in terms of digital sovereignty. Notably, all stakeholder needs and desires are considered in elicitation by identifying the stakeholder representatives of each stakeholder type who are then involved in elicitation activities. Finally, the requirement register is cross-checked against the stakeholder register to validate the basic requirements that have already been elicited and do not need to be refined or modified further.

Step 3 document analysis Documents are analysed to learn about the characteristics, needs, desires, and requirements of various stakeholder types. Document analysis is specifically designed to uncover missing basic requirements, enabling the growing collection of quality requirements to be realised in the future course of MERDigS. Initially, in the analysis, informative documents are found, selected, and saved in a folder. Thus, subsequent document analyses can be conducted systematically and purposefully [32]. Furthermore, with subsequent document analysis, ambiguous statements that are considered indicators of noncommunicable requirements can be further discussed in subsequent elicitation activities [22, 33, 34]. Additionally, a glossary is established for consistent documentation and understandable communication as early as possible to reduce effort due to inconsistencies and ambiguities [13]. Finally, potential requirements are derived and justified based on the information about the needs and desires of stakeholder types.

Step 4 Focus group-like workshops This step entails holding focus group-style workshops with four to nine stakeholder representatives, divided into two parts. The first part of each workshop is primarily focused on validating elicited basic requirements, after which the basic requirements no longer need to be discussed in subsequent interviews because they have either been validated or (provisionally) transformed into quality requirements. This is conducted through yes/no questions answered by the focus group team members [26]. A dichotomous answer format is suggested because of the expected simplicity of finding an answer. The simplicity of the answer selection results from the focus group team is expected to unambiguously agree with the basic requirements. The second part of each workshop is inspired by a collaborative elicitation technique termed the KJ method [13, 35]. Individual brainstorming regarding new requirements is conducted in this section, followed by a discussion of the requirements. Individual brainstorming ensures the elicitation of requirements, after which the elicited requirements can be justified, discussed, and qualitatively improved within the focus group team [36]. Subsequently, use-case considerations are recommended. Thus, the use cases of the project are refined through activities that are discussed step-by-step with the focus group team. Potential constraints and enablers of the digital sovereignty of the considered stakeholder group are perceived and discussed in an intuitive and application-oriented manner during this process [19, 37]. Following the workshops, within the meaning of introspection, the members of the focus group team are provided with the opportunity to add previously unmentioned requirements [13]. These may only emerge through conscious awareness and observation of everyday behaviour [14]. For the same reason, the workshop invitation already includes impulses on the topic of digital sovereignty, which should be perceived more consciously in everyday life before the workshops.

Step 5 Conducting interviews Among other things, the information gathered in steps 1 to 4 is used to conduct the interviews in step 5 to the smooth extent possible and maximise the benefits of direct communication with stakeholder representatives. Therefore, all pertinent information gathered thus far is compiled in an expanded subject area register, which is used to design and structure questions for interviews with stakeholder representatives in the interview guidelines. The interview guidelines contain questions for each subject area, to which answer options and follow-up questions are assigned, creating a tree-like structure [22, 36]. Herein, the initial questions are primarily based on the needs and desires of stakeholders, and the follow-up questions are primarily based on the (quality) requirements in the field of digital sovereignty [18]. The interviews are semi-structured, enabling and encouraging spontaneous questions that capture the cognition of stakeholder representatives [38]. Other parts of the interviews are used to clarify ambiguous statements, open-ended questions, and potential requirement conflicts, all of which can be used to elicit noncommunicable quality requirements. These results are shown in Table 2.

4 Evaluation

The plausibility of MERDigS's approach was evaluated through three expert interviews with experienced project members from the state-funded project 'SDIKA–Schaufenster Sichere Digitale Identitäten Karlsruhe' (Showcase Secure Digital Identities Karlsruhe). Project members from SDIKA represent both potential applicants and members of the affected stakeholder groups in MERDigS. Thus, they analysed MERDigS from two different perspectives, which is why their background knowledge was particularly useful in evaluating MERDigS. Experts can be assigned to public administration stakeholder groups (expert 1), scientific organisations (expert 2), and business organisations (expert 3). SDIKA is another software development project for empowering citizens and organisations in the state of digital sovereignty. Consequently, project members are assumed to have specialised knowledge and critical thinking skills in the field of digital sovereignty.

Conducting expert interviews consumed more than 4,5 h of video material. The preparation of results of transcribed expert interviews was based on specific upper and lower categories. The upper categories comprised three categories: 'Meeting the requirements', 'Quality of method application', and 'Quality of method results', to which different evaluation criteria were subordinated.

In summary, experts assessed the MERDigS approach as plausible and meaningful. According to Expert 1, many useful modules were also used in practice, and high-value artefacts were generated, justifying that the use of MERDigS would be extremely useful for the organisation. Expert 2 examined the incorporation of every single activity, including the incorporation of optional modules, as scientifically justified. In their overall picture, they lead to the 'perfection' of MERDigS. Nevertheless, a demand for additional degrees of freedom in MERDigS arises. This demand can be traced back to the heterogeneity of software development projects. As the structure of MERDigS is extremely complex, Expert 3 seems particularly harrowed to MERDigS.

Table 1 Subject areas in digital sovereignty to elicit requirements based on [10, 25, 26, 28]

Subject area	Description	Ancillary subject areas
Company and business secrets	Addresses the question of who may disclose organisation-specific information arising from commercial or technical spheres and represents significant corporate assets [29] in an authorised manner	Authorisation and information confidentiality
Competence deployment	Includes the required and desired use of competencies and knowledge of the stakeholder using the software	Basic knowledge, autonomy, and assistance
Control	Addresses the method through which and the extent to which the stakeholder has control over the software and its processes	Control delivery and responsibility
Decision-maker	Concerns the question of the areas in which the software can be used and take the decisions from humans	Responsibility and substitution
Flexibility	Enquires about the possibilities that the software offers to extend the software independently and the stakeholder to adapt it	Scalability, networkability, and adaptability
Functionality	Includes the completeness regarding the software functions and the expected functional scope by the stakeholders	Adequacy and correctness
Health	Enquires about the influence of the software on mental and physical health and methods through which these aspects can be positively influenced by the software	Communication, extent and addiction, and relief
Identity	Is concerned with the management, storage, protection, and authenticity of identity data	Identity management, authorisation, authenticity, and protection of data privacy
Independence	Comprises the degree of dependency of the software provider on the software stakeholders operating and using the software	Control, trust, and foreign companies
Infrastructure	Enquires about the basis on which the software should be built	Origin of infrastructure, integrity, and security
Interoperability	Includes the compatibility of the software with existing digital technologies, facilitating switching between different technology providers	Interconnectivity, integrity, accessibility, portability, and changeability

(continued)

Table 1 (continued)

Subject area	Description	Ancillary subject areas
IT security	Addresses the procedure through which and the extent to which the stakeholder expect to be protected from threats and external attacks	Resilience, freedom from manipulation, stability, traceability, and delegation of rights
Neutrality	Enquires to what extent the software should remain neutral from laws and restrictions and to what degree it should influence the stakeholder	Influence and Limitations
Performance	Includes the ability of the software to serve software stakeholders so that its use results in a benefit to comparable software	Timing, cost, benefit, effectiveness, and efficiency
Platforms	Deals with the question of the procedure through which and by whom the platforms should be designed	Market fragmentation and trust
Privacy	Deals with the release of the identity of the stakeholder and the identities surrounding the stakeholder/the possibility of using the software without inference to the stakeholder's identity	Nonconnectivity, communication, and identity
Reliability	Describes the ability of the software to maintain a level of performance under certain conditions over a certain period	Fault tolerance and durability
Sustainability and maintenance	Enquires the extent to which sustainable resources and a modular structure are necessary	Modularity, resources, and reusability
Usability	Includes the ability of software to be understood, learnable, and executable	Intelligibility, comprehensibility, simplicity, learnability, and attractiveness
Self-fulfilment	Is concerned with benefits derived by stakeholders from the use of the software compared with the real world and to what extent the software can contribute to the realisation of interests and personality development	Digital presence and degrees of freedom
Transparency	Comprises the level of abstraction and the extent to which a software discloses its algorithmic decision-making processes, usage implications, and background processes to its stakeholders	Information, understandability, openness, and traceability
Trust	Deals with trust granting and the elimination of uncertainties	Security, protection, and control

Table 2 Supplementary interview modules

Interview module [source]	Description
Engagement scenarios [24, 39]	In this module, the stakeholder representative is presented with predefined scenarios about which questioning is performed. The scenarios are intended to broaden the stakeholder representative's awareness of the project to recognise the advantages and disadvantages, as well as the context of use, and can add requests
Short Stories [18, 39]	In this module, stakeholder representatives narrate a story regarding their expectation from the software development project. The story should include a goal, a description, affected stakeholders, limitations, and alternatives. From the description, information about what the stakeholder representative wants, why he wants it, when he wants it, and where he wants it, can be derived. The facilitator can then ask follow-up questions about the story
Presenting response options [33]	If the response options to a question in the interviews are highly likely to be completed, the response options could be presented to the stakeholder representative. Therefore, the stakeholder representative must decide on an answer option, agree to the answer option in numbers, and state the reasons for selecting the response option
Rankings [40, 41]	If inconsistencies are observed in the requirements for each topic and stakeholder type, the requirements can be ordered by the stakeholder representative according to their preferences and then justified. The requirements can be written on cards for this purpose
Statements [34]	Here, an ambiguous statement is presented to the stakeholder representative as a statement indicating a specific usage objective. First, the stakeholder representative explains the statement. Then, advantages and disadvantages that support the explanation are added. If an ambiguous statement is presented to all stakeholder representatives in the interviews as a statement, the understanding that receives the most relevance and agreement is the unfolded understanding of an initial ambiguous statement

The complete execution of MERDigS may be extremely resource-intensive for smaller software development projects. Therefore, more optional modules and focus points should be set in the process model. Existing tasks such as establishing the glossary and existing steps such as step six or new activities such as distributing the survey results could therefore be incorporated as 'optional' tasks, activities, or steps. Software development projects should then decide whether these tasks either lead to excess or necessary additional work. As Expert 1 notices a tendency to incorporate additional feedback loops, large software development projects, such as government-subsidised projects, might notice necessary additional effort in these tasks.

In the following, the evaluation is presented in detail, according to the three upper categories.

4.1 Meeting the Requirements

Adaptation. A target criterion of MERDigS is that it should comprise existing published methods, procedures, techniques, tools, and languages. Whether this criterion is met was not explicitly asked in the expert interviews. Notably, implicit comments suggest that MERDigS is built on a stable theoretical foundation based on high-quality and comprehensive literature research. Regarding the research methodology, Expert 1 notes that the complete construction kit of requirements engineering was used and MERDigS is completely mature and without gaps. In Addition, Expert 3 states that MERDigS uses various methods and compares many different theories.

Completeness. The target criterion of completeness includes completeness regarding the dimensions of digital sovereignty and the types of requirements. A fulfilment of the completeness regarding the dimensions of digital sovereignty is supported by the provision of a subject area list as a tool for the creation of the subject area register. However, according to Expert 2, it depends significantly on the applicant of MERDigS and the intrinsic incentives of the applicant. Finally, users must set up the subject area register and adapt it to their software development project. With the (superficial) questioning in expert interviews, the extent to which noncommunicable requirements are also collected has not been elucidated. Moreover, this raises the question of the usefulness of interventions for eliciting noncommunicable requirements. For efficiency reasons, smaller projects might have to skip activities, such as recording ambiguous statements, defined for this purpose.

Applicability. The applicability criterion includes the applicability of MERDigS to all stakeholders of digital sovereignty. All the experts agreed with the grouping, which is necessary to fulfil this criterion, from Chap. 3.1. In addition, after a brief reflection period, the inclusion of machines, which should also be considered currently, to stakeholder groups was suggested by Expert 1.

4.2 Quality of Method Application

To evaluate the quality of MERDigS' application, three criteria of *empirical feasibility, acceptability,* and timeline were considered together per step. These criteria describe the smooth implementation of MERDigS on all stakeholders, the appropriateness and usefulness of MERDigS, and the adherence to a defined time frame. To this end, the experts were asked whether the individual steps tend to run smoothly, haltingly, or with increased complexity.

Step one is described as a smooth step that sounds good and also necessary. Notably, the creation of the topic area register is suitable for building up a broader understanding of the topic of digital sovereignty. For each of the experts, the stakeholder analysis in step two is a critical step being appropriately inserted at this point. In step three, the experts initially agree that document analysis is useful. However, they disagree on the appropriateness of the complexity of individual activities. In particular, the creation of a glossary is disputed, which according to Expert 1 should be defined even more broadly, in the sense of an 'expanded language model'. Experts 2 and 3 consider the creation of a glossary to be unnecessary in practice (but not in theory) as glossaries are of no use afterwards. Step four also demands further degrees of freedom and more flexibility. For example, Expert 1 demands further, optional feedback loops with project members in this step, whereas Expert 2 would omit it to shorten the elicitation process. Simultaneously, he suggests using MERDigS to define the goals of the focus group-like workshop without explaining the achieving of these goals. The approaches of suggesting different tools as options and orienting the individual brainstorming task towards subject areas in MERDigS identify the evaluation as critical. In addition, Expert 3 perceives the step as extremely noticeable, well thought out, and meaningful. In step five, all the experts agree that the step runs smoothly and is important. Expert 1 again recommends extending the step to include an optional feedback loop. Step six initially runs smoothly according to Experts 1 and 3. In practice, however, according to Expert 2, questionnaires are most likely to be omitted, especially in small projects. Expert 3 backs this up by noting the effort involved in finding a representative sample.

The fulfilment of the empirical feasibility criterion is also proven by the consensus of the experts that MERDigS is learnable. In addition, according to all the experts, the time required to conduct MERDigS is justified, especially by the relevance and actuality of the topic of digital sovereignty. Moreover, requirements elicitation is generally worthwhile, as the corresponding time amortises over time.

4.3 Quality of Method Results

Understandability and structure. These criteria include a noticeable, logical, and complete structure of the generated artefacts. To evaluate these criteria, a drafted requirements document was presented to the experts. Thereupon, all the experts agreed that it is generally understandable and well-structured. Expert 1 validated the content of the requirements document by noticing that similar points were found in his requirements document. To obtain an additional structure in the requirements document, Expert 1 suggests adding standardised diagrams of the individual process steps. Expert 2 mentions the possibility of structuring the requirements in the requirements document using the subject areas.

Semantic correctness. Semantic correctness indicates whether the true requirements of the stakeholders in the area of digital sovereignty can be represented by the collected requirements without errors or losses [42]. Given that the guidelines of MERDigS, in which various cross-checks are included, are followed correctly, the experts estimate the potential for errors to be low and manageable. However, the risk cannot be completely excluded, as the semantic correctness is strongly dependent on the applicant. Expert 1 believes that repeated interviews with different stakeholder representatives provide additional security for obtaining correct results. Finally, all the experts believe that enough tasks are incorporated to avoid inconsistencies, conflicts, and contradictions.

Relevance. The criterion of relevance encompasses the scope of the requirements raised, which should not fall outside the topic of digital sovereignty. To fulfil this criterion, the area of digital sovereignty must be correctly delimited. Through oral questioning, this criterion is crucial to evaluation, although Expert 2 dares to claim that the subject area register is used to collect requirements in a defined area. These specify the area of digital sovereignty and the subject areas in which requirements are to be collected, as some are deliberately excluded at the beginning. Simultaneously, Expert 2 pointed out the risk of excluding individual subject areas, as the time when all requirements for a subject area will be collected is unknown.

5 Conclusion and Future Work

Digital sovereignty is a growing desire of various stakeholders, whose requirements vary significantly depending on both the specific context and stakeholder. Studies into these needs are currently conducted to a greater extent. Requirements elicitation is not yet sufficiently methodologically supported. Existing methods can be used to elicit some requirements in the area of digital sovereignty. However, these rapidly run into contradictions with other unrealised requirements and lead to negative feedback. MERDigS, a method for eliciting requirements in the area of digital sovereignty, provides a solution for this. The few existing scientific publications with direct reference to digital sovereignty and adaptable methods were included in the development of MERDigS.

The combination of various complementary techniques enables a progressively stronger focus on (quality) requirements arising specifically from the domain of digital sovereignty. The increasing focus allows stakeholders to increasingly respond to their needs and targeted desires, ultimately facilitating the implementation of the human-centred approach of MERDigS by conducting interviews. However, as stakeholders are only involved in the last two steps, the additional effort for implementation is manageable. Furthermore, with the implementation of the subject area register, MERDigS faces the challenge of collecting requirements in the area of digital sovereignty and thus in a difficult-to-understand framework. The subject area register specifies the subject areas in which requirements must be collected, thereby meaningfully limiting the area of digital sovereignty. Moreover, the process model of MERDigS is determined by specifications, recommendations, and optional modules. The allowed degrees of freedom indicate that MERDigS can be applied to different software development projects.

Nonetheless, MERDigS should be more responsive to project heterogeneity and thus be built more generically. Finally, expert interviews indicate that the specifications in MERDigS are overly detailed for small software development projects and underdeveloped for large software development projects.

Furthermore, additional efforts are required to empower stakeholders in the state of digital sovereignty. This motivates the extension of the set framework of MERDigS. Finally, although the requirements are listed with the application of MERDigS, their technical implementation has not been elucidated. Thus, the framework could be expanded to include a direct link to the developers of the software development project, enabling technical implementation options to be discussed concurrently with requirements elicitation. Another reason for extending the set framework of MERDigS is the ongoing conflict potential of requirements in the area of digital sovereignty. Above all, this potential conflict exists among different stakeholder groups that pursue and assume divergent goals and functions. Therefore, in an extension of MERDigS, it can be negotiated as to whose requirements are dominant in which subject area.

As the requirements in the area of digital sovereignty have to be negotiated both with and between the various stakeholders, no state of digital sovereignty can satisfy everyone to the respective individual maximum. To guarantee the state's digital sovereignty to many stakeholders, conscious joint thinking and action against the dominance of a few leading digital companies with solely capital-oriented decision-making structures are necessary. Consequently, collaborative solutions should be developed to enable some dependence to demonstrate the willingness and trust of those sharing the same values and goals. To this end, digital barriers should be removed and digital competencies be developed.

Acknowledgements The content of this paper is result of the project 'SDIKA—Schaufenster Sichere Digitale Identitäten Karlsruhe' (Showcase Secure Digital Identities Karlsruhe). The goal of the project is to use digital identities to connect people, organisations, and processes. The values of digital sovereignty, fairness, and interoperability are guiding principles of the project and for the regional showcase. This project is supported by the Federal Ministry for Economic Affairs and Climate Action (BMWK) on the basis of a decision by the German Bundestag.

References

1. Couture S, Toupin S (2019) What does the notion of "sovereignty" mean when referring to the digital? New Media Soc 21:2305–2322. https://doi.org/10.1177/1461444819865984
2. Creemers R (2020) China's conception of cyber sovereignty: rhetoric and realization. SSRN Journal. https://doi.org/10.2139/ssrn.3532421
3. Pohle J (2020) Digitale Souveränität: Ein neues digitalpolitisches Schlüsselkonzept in Deutschland und Europa. Berlin
4. Stubbe J, Schaat S, Ehrenberg-Silies S (2019) Digital souverän?: Kompetenzen für ein selbstbestimmtes Leben im Alter. Bertelsmann Stiftung, Gütersloh
5. Krupka D, Kranich L, Schipanksi T, Bending T, Steinacker K, Zimmermann J et al (2020) Schlüsselaspekte Digitaler Souveränität. Berlin
6. Desmarais-Tremblay MWH (2020) Hutt and the conceptualization of consumers' sovereignty. Oxf Econ Pap 72:1050–1071. https://doi.org/10.1093/oep/gpaa015
7. Ernst C (2020) Der Grundsatz digitaler Souveränität: Eine Untersuchung zur Zulässigkeit des Einbindens privater IT-Dienstleister in die Aufgabenwahrnehmung der öffentlichen Verwaltung. Duncker & Humblot, Berlin

8. Floridi L (2020) The fight for digital sovereignty: what it is, and why it matters especially for the EU. Philos Technol 33:369–378. https://doi.org/10.1007/s13347-020-00423-6
9. Ahmed S, Kanwal HT (2014) Visualization based tools for software requirement elicitation. In: 2014 international conference on open source systems and technologies (ICOSST), 18.12.2014–20.12.2014. IEEE, Lahore, Pakistan, pp 156–159. https://doi.org/10.1109/ICOSST.2014.7029337
10. Lim T-Y, Chua F-F, Tajuddin BB (2018) Elicitation techniques for internet of things applications requirements. In: Unknown (ed) The 2018 VII international conference; 14.12.2018–16.12.2018. ACM Press, Taipei City, Taiwan, New York 2018. pp 182–188. https://doi.org/10.1145/3301326.3301360
11. Al-Zawahreh H, Almakadmeh K (2015) Procedural model of requirements elicitation techniques. In: Boubiche DE, Hidoussi F, Cruz HT (ed) IPAC '15: International conference on intelligent information processing, security and advanced communication, Batna Algeria. https://doi.org/10.1145/2816839.2816902
12. Weinreuter M (2022) Methode zur Entwicklung von Anforderungen im Bereich der digitalen Souveränität: Karlsruher Institut für Technologie (KIT). https://doi.org/10.5445/IR/1000151128
13. Angelis G, Ferrari A, Gnesi S, Polini A (2016) Collaborative requirements elicitation in a European research project. In: Ossowski S (ed) SAC 2016: symposium on applied computing. Pisa Italy, 1282–1289. https://doi.org/10.1145/2851613.2851760
14. Tiwari S, Rathore S (2017) A methodology for the selection of requirement elicitation techniques. https://doi.org/10.48550/arXiv.1709.08481
15. Ferrari A, Spoletini P, Gnesi S (2016) Ambiguity and tacit knowledge in requirements elicitation interviews. Requirements Eng 21:333–355. https://doi.org/10.1007/s00766-016-0249-3
16. Sutcliffe A, Sawyer P (2013) Requirements elicitation: Towards the unknown unknowns. In: 2013 IEEE 21st international requirements engineering conference (RE); 15.07.2013–19.07.2013. IEEE, Rio de Janeiro-RJ, Brazil. pp 92–104. https://doi.org/10.1109/RE.2013.6636709
17. Umber A, Naweed MS, Bashir T, Bajwa IS (2012) Requirements elicitation methods. Adv Mater Res 433:6000–6606. https://doi.org/10.4028/www.scientific.net/AMR.433-440.6000
18. Ali N, Lai R (2017) A method of requirements elicitation and analysis for global software development. J Softw Evol Process. https://doi.org/10.1002/smr.1830
19. Thew S, Sutcliffe A (2018) Value-based requirements engineering: method and experience. Requirements Eng 23:443–464. https://doi.org/10.1007/s00766-017-0273-y
20. Gupta AK, Deraman A (2019) Algorithmic solution for effective selection of elicitation techniques. In: 2019 International conference on computer and information sciences (ICCIS). https://doi.org/10.1109/ICCISci.2019.8716378
21. Atukorala NL, Chang CK, Oyama K (2016) Situation-Oriented requirements elicitation. In: IEEE 40th annual computer software and applications conference (COMPSAC). IEEE, Atlanta, USA, pp 233–238. https://doi.org/10.1109/COMPSAC.2016.191
22. Zhi Q, Zhou Z, Morisaki S, Yamamoto S (2019) An approach for requirements elicitation using goal, question, and answer. In: 2019 8th international congress on advanced applied informatics (IIAI-AAI). IEEE, Toyama, Japan. pp 847–852. https://doi.org/10.1109/IIAI-AAI.2019.00172
23. Neetu KS, Pillai AS (2014) A study on project scope as a requirements elicitation issue. In: International conference on computing for sustainable global development (INDIACom). IEEE, New Delhi, India, pp 510–514. https://doi.org/10.1109/IndiaCom.2014.6828190
24. Vujicic T, Scepanovic S, Jovanovic J (2016) Requirements elicitation in culturally and technologically diverse settings. In: 5th Mediterranean conference on embedded computing (MECO). IEEE, Bar, Montenegro, pp 464–467. https://doi.org/10.1109/MECO.2016.7525693
25. García-López D, Segura-Morales M, Loza-Aguirre E (2020) Improving the quality and quantity of functional and non-functional requirements obtained during requirements elicitation stage for the development of e-commerce mobile applications: an alternative reference process model. IET software 14:148–58. https://doi.org/10.1049/iet-sen.2018.5443

26. Silva A, Pinheiro PR, Albuquerque A, Barroso J (2017) Evaluation of an approach to define elicitation guides of non-functional requirements. IET Software 221–228. https://doi.org/10.1049/iet-sen.2016.0302
27. Burnay C (2016) Are stakeholders the only source of information for requirements engineers? Toward a taxonomy of elicitation information sources. ACM Trans Manage Inf Syst 7. https://doi.org/10.1145/2965085
28. Ferraris D, Fernandez-Gago C (2020) TrUStAPIS: a trust requirements elicitation method for IoT. Int J Inf Secur. https://doi.org/10.1007/s10207-019-00438-x
29. Alpers S (2019) Modellbasierte Entscheidungsunterstützung für Vertraulichkeit und Datenschutz in Geschäftsprozessen. KIT Scientific Publishing. https://doi.org/10.5445/KSP/1000094545
30. Palomares C, Franch X, Quer C, Chatzipetrou P, López L, Gorschek T (2021) The state-of-practice in requirements elicitation: an extended interview study at 12 companies. Requirements Eng 26:273–299. https://doi.org/10.1007/s00766-020-00345-x
31. Ryan MJ (2014) The role of stakeholders in requirements elicitation. INCOSE Int Symp 24:16–26. https://doi.org/10.1002/j.2334-5837.2014.tb03131.x
32. Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering. 2nd ed. Keele University
33. Al-Alshaikh HA, Mirza AA, Alsalamah HA (2020) Extended rationale-based model for tacit knowledge elicitation in requirements elicitation context. IEEE Access. https://doi.org/10.1109/ACCESS.2020.2982837
34. Anwar H, Khan SUR, Iqbal J, Akhunzada A (2022) A tacit-knowledge-based requirements elicitation model supporting COVID-19 Context. IEEE Access 10. https://doi.org/10.1109/ACCESS.2022.3153678
35. Scupin R (1997) The KJ method: a technique for analyzing data derived from Japanese ethnology. Hum Organ 56:233–237. https://doi.org/10.17730/humo.56.2.x335923511444655
36. Farinha C, Da Mira Silva M (2013) Requirements elicitation with focus groups: lessons learnt
37. Rocha ST, Winckler M, Bach C (2020) Evaluating the usage of predefined interactive behaviors for writing user stories: an empirical study with potential product owners. Cogn Tech Work 22:437–457. https://doi.org/10.1007/s10111-019-00566-3
38. Kanwal A (2019) Requirements engineering: elicitation techniques. Int J Sci Eng Res 10:154–162
39. Wahbeh A, Sarnikar S, El-Gayar O (2020) A socio-technical-based process for questionnaire development in requirements elicitation via interviews. Requirements Eng 25:295–315. https://doi.org/10.1007/s00766-019-00324-x
40. Renzel D, Behrendt M, Klamma R, Jarke M (2013) Requirements Bazaar: social requirements engineering for community-driven innovation. In: 2013 IEEE 21st international requirements engineering conference (RE), 15.07.2013–19.07.2013. IEEE, Rio de Janeiro-RJ, Brazil. pp 326–327. https://doi.org/10.1109/RE.2013.6636738
41. Mukherjee N, Zabala A, Huge J, Nyumba TO, Adem EB, Sutherland WJ (2018) Comparison of techniques for eliciting views and judgements in decision making. Methods Ecol Evol 9:54–63. https://doi.org/10.1111/2041-210X.12940
42. Al-Subaie H, Maibaum T (2006) Evaluating the effectiveness of a goal-oriented requirements engineering method, pp 8–19 (2006). https://doi.org/10.1109/CERE.2006.3

A Hybrid Federated Learning-Based Ensemble Approach for Lung Disease Diagnosis Leveraging Fusion of SWIN Transformer and CNN

Asif Hasan Chowdhury, Md. Fahim Islam, M. Ragib Anjum Riad, Faiyaz Bin Hashem, Md Tanzim Reza, and Md. Golam Rabiul Alam

Abstract The significant advancements in computational power create a vast opportunity for using artificial intelligence in different applications of healthcare and medical science. **A hybrid FL-enabled ensemble approach for lung disease diagnosis leveraging a combination of SWIN transformer and CNN** is the combination of cutting-edge technology of AI and federated learning. Since, medical specialists and hospitals will have shared data space, based on that data, with the help of artificial intelligence and integration of federated learning, we can introduce a secure and distributed system for medical data processing and create an efficient and reliable system. The proposed hybrid model enables the detection of COVID-19 and pneumonia based on X-ray reports. We will use advanced and the latest available technology offered by TensorFlow and Keras along with Microsoft-developed vision transformer that can help to fight against the pandemic that the world has to fight together as a united. We focused on using the latest available CNN models (DenseNet201, InceptionV3, VGG19) and the transformer model SWIN transformer in order to prepare our hybrid model that can provide a reliable solution as a helping hand for the physician in the medical field. In this research, we will discuss how the federated learning-based hybrid AI model

A. H. Chowdhury (✉) · Md. F. Islam · M. R. A. Riad · F. B. Hashem · M. T. Reza · Md. Golam Rabiul Alam
BRAC University, Dhaka 1212, Bangladesh
e-mail: asif.hasan.chowdhury@g.bracu.ac.bd

Md. F. Islam
e-mail: md.fahim.islam@g.bracu.ac.bd

M. R. A. Riad
e-mail: m.ragib.anjum.riad@g.bracu.ac.bd

F. B. Hashem
e-mail: faiyaz.bin.hashem@g.bracu.ac.bd

M. T. Reza
e-mail: tanzim.reza@bracu.ac.bd

Md. Golam Rabiul Alam
e-mail: rabiul.alam@bracu.ac.bd

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_77

can improve the accuracy of disease diagnosis and severity prediction of a patient using the real-time continual learning approach and how the integration of federated learning can ensure hybrid model security and keep the authenticity of the information.

Keywords AI · VGG19 · InceptionV3 · DenseNet201 · SWIN transformer · Federated learning · Privacy

1 Introduction

Integrating artificial intelligence with medical science has created a new dimension to the treatment world. Computer-assisted diagnosis can help doctors to sense any forthcoming lethal diseases beforehand. Nowadays, doctors across the world tend to rely more on AI as it is improving swiftly. We are looking to develop a system that can identify lung diseases that can help medical people during the treatment procedure. We are aware that we need to be very cautious to develop a system that will analyze patients' medical reports and identify the disease of patients. We are focusing to use the AI-driven approach to address the lung disease of patients. In this research, we proposed an ensemble method to detect lung diseases. We focus to achieve preferable accuracy with better performance; therefore, we build an ensemble method for our research. We have ensembled several latest AI-based algorithms like VGG19, InceptionV3, DenseNet201, and vision transformer developed by Microsoft. We have combined the outcome from this algorithm to develop a model that will be unique and reliable for lung disease detection. Furthermore, we took the help of federated learning to ensure the data privacy of sensitive patient medical images. We want to build a network through federated learning where different hospitals will stay connected together and share their effective treatment models. These effective models will be used to improve the performance of the central model which will be considered the core of the entire system. This global model will be updated based on the outcome from the local model through the federated learning-based network to ensure high security during the weight transfer process between the models that will be in different parts of the world.

1.1 Research Objective

The main objective of this research is to build a fusion model using transfer learning and a transformer model to save the patient from ARDS, a severe state of the lung. We aim to improve the outcomes of studies on existing transfer learning models by adding SWIN transformers to make a fusion model, and also using federated learning, we aim to ensure healthcare data security, low latency, and less power consumption.

- Lung disease detection using the deep CNN model to analyze the severe conditions of patients.
- Improve existing transfer learning models by building a new fusion model that combines transfer learning and transformer learning.
- Integration of shifted window (SWIN) transformer model for better accuracy and detection.
- Utilization of federated learning models for ensuring data security, low latency, less power consumption, and better accuracy.

2 Literature Review

Kassania et al. proposed a deep CNN approach to detect COVID-19 from X-ray and CT images [1]. The authors tried to get a better solution to the over-fitting issue in deep learning due to the small number of training images by using a transfer learning strategy. Firstly, Kassania et al. applied the image normalization technique to get better visual quality of input images. In the feature extraction step, the authors used a transfer learning strategy to lessen computational resources and accelerate the convergence of the network as their dataset is very limited. Finally, the authors developed a Web-based application to assist doctors in detecting COVID-19 by uploading X-ray or CT images. This research contains some limitations such as few training data samples and security assurance. In our paper, we fed our model with a comparatively larger training set while ensuring the privacy of patients. We have developed a fusion model to get more accuracy in an efficient way.

Hemdan et al. introduced a framework of deep learning, named COVIDX-Net to diagnose COVID-19 from X-ray images [2]. The COVIDX-Net framework consists of seven DCNNs of different architectures, and those are VGG19, DenseNet201, InceptionV3, ResNetV2, InceptionResNetV2, Xception, and MobileNetV2. The authors fed their models with a limited number of data. Hemdan et al. got better results using VGG19 and DenseNet201, whereas the result with InceptionV3 was not satisfactory. This paper shows a comparison among existing DCNN models with limited data.

Minaee et al. [3] trained 4 state-of-the-art convolutional networks for COVID-19 detection. Jiang et al. proposed a model which is a combination of SWIN transformer and transformer and transformer to make a classification of COVID-19 from a dataset of X-ray images [4]. The existing conventional models have slow computational power and large sizes. Using the SWIN transformer model can increase the computational speed with the size of the image. In this paper, Gu et al. proposed a fusion model that combines SWIN transformer blocks and a lightweight U-Net type model that has an encoder–decoder structure [5].

Li et al. have discussed the mechanism of federated learning for securing data and overcoming challenges [6]. First, the authors have mentioned the challenges, one might face implementing federated learning such as expensive communication, system heterogeneity, statistical heterogeneity, and privacy concerns. Then, Li et al.

come up with solutions to those challenges. For more communication efficiency, the authors have pointed to multiple methods like local updating and decentralized training. As model training may differ for devices' hardware specifications, the authors recommend using an asynchronous scheme that applies an optimization algorithm parallelly. Sometimes, data for federated learning are divided among devices non-identically, to resolve this problem, meta-learning and multitask learning have been added to federated settings [7].

A framework, named MOCHA [8], is used for learning separate but related models for each device. Bonawitz et al. [9] introduced a protocol for individual model updates. In this protocol, the central model will not be able to see the local updates but will observe the aggregated result at the end of every round. This method is an inspiration for our proposed model where we will implement federated learning to overcome privacy leakage.

3 Methodology

In Fig. 1, we have shown the top-level overview of the proposed lung disease detection system. First, we will collect the data. Next, we will start to train the existing transfer-based learning models using our dataset. We have planned to work with VGG19, InceptionV3, and DenseNet201 model. Then, we will save the best-trained model individually and ensemble them together. Next, we will start to train the transformer-based model that we decided to work with the SWIN transformer model. We will

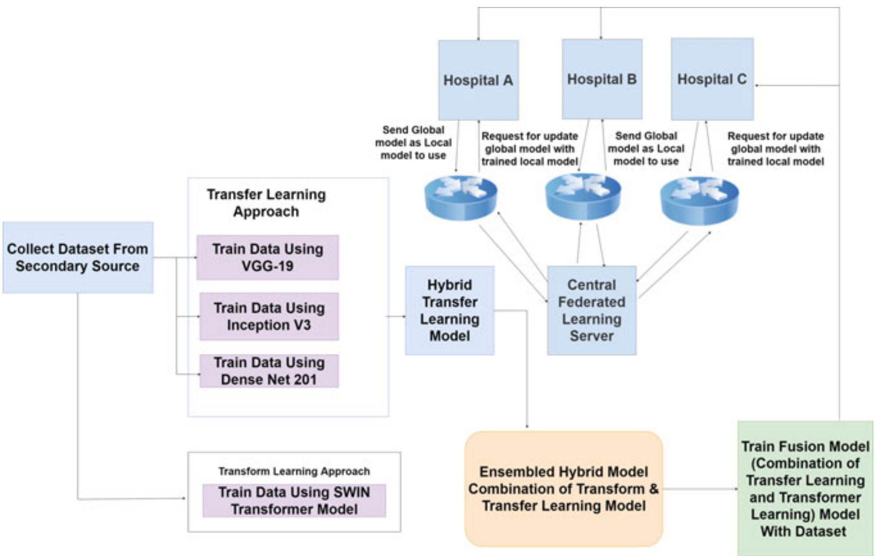


Fig. 1 Top level over view of proposed disease detection system

also train this model using our dataset. Next, we will combine the trained SWIN transformer model with the transfer learning-based ensemble model to create our own hybrid model. Next, we will train the hybrid model with our available dataset to complete our model training and validation. Moreover, we are using a federated learning approach to secure each individual model that will be held by the hospitals. Hospitals will share the best finding outcome with the global server to ensure better accuracy and outcome.

3.1 Dataset Collection Process

Data collection is the most significant task to start building the CNN model. The initial step of the work plan is to collect data from different primary sources. We know medical data is sensitive and difficult to manage. We initially looked to find medical data from different hospitals. In most cases, we were not able to manage the same disease-related information. Then, we looked at different disease-related papers for the dataset. We get different datasets, but still being open-source datasets, some people alternated the large dataset with wrong files and corrupted files. Thus, we have become careful enough during the selection process of the dataset (Fig. 2).

3.2 Dataset Analysis and Interpretation

we processed the image data using the available pre-processing techniques. We have identified corrupted images in the first place. Then, we removed the wrong image data. For example, the X-ray dataset contains CT-image data. Then, we figure out the

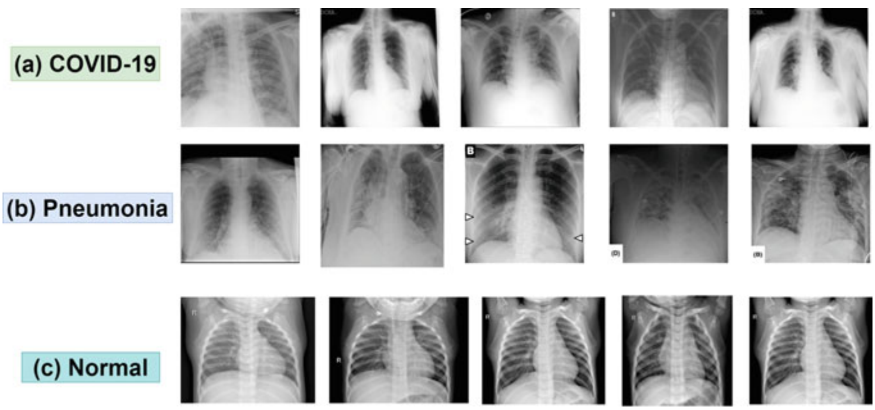


Fig. 2 Sample dataset of **a** COVID-19; **b** Pneumonia; **c** Normal

number of images available for the training process and increase the data by the data augmentation process. Next, we used re-scaling, mage-rotation, horizontal-rotation, and zoom-range for the image processing process. We need to split the dataset into two categories known as training sets and testing sets. The train set will do the task of training the dataset and preparing local models for different hospitals. The test set will do the job of testing the predicted diseases. We followed the convention of the training set(80%) and testing set(20%). We will have more accurate results if we can increase the ratio of the training set.

3.3 Classification and Decision Classifier

As our paper requires multiple predictions, we implement the VGG19, InceptionV3, and DenseNet 201 for it. In addition, we have used a transformer-based learning model that is the SWIN transformer model. These models will exhibit real-time predictions for each individual. The probabilistic results of the model will help patients and medical practitioners to detect the disease.

VGG19 VGG19 is the latest version of the visual geometry group model series. This model series is the successor of the AlexNet. This model consists of 19 layers. Out of 19 layers, 16 layers are convolutional layers, 3 fully connected layers, 5 MaxPool layers, and 1 softmax layer.

In order to categorize the photos into 1000 object categories, Simonyan and Zisserman (2014) presented the VGG19. There are numerous 3×3 filters used by each convolutional layer. Because each convolutional layer uses numerous 3×3 filters, it is a highly well-liked technique for classifying images.

InceptionV3 The third generation of inception convolutional neural network designs is known as InceptionV3. Among other improvements, the InceptionV3 convolutional neural network architecture makes use of label smoothing, factorized 7×7 convolutions, and the addition of an auxiliary classifier to move label information lower down the network (along with the use of batch normalization for layers in the sidehead).

The inception architecture is built to function successfully even when memory and compute resources are severely limited. Though the architectural simplicity of VGGNet is appealing, it comes with a considerable computational cost when assessing the network. As inception is lower while higher-performing successors, it is feasible to utilize the e inception networks in big-data scenarios.

The architecture of an InceptionV3 is progressively built, step-by-step. 1. Factorized convolutions: This decreases the number of parameters used in a network, which lowers computational efficiency. It also monitors the effectiveness of the system. 2. Smaller convolutions: This causes training to go more quickly by substituting smaller convolutions for larger ones. Say a 5×5 filter has 25 parameters; replacing it with two 3×3 filters results in only 18 ($3 * 3 + 3 * 3$) parameters [10].

DenseNet201 A typical convolutional neural network is started with an input image and runs through the network to get a predicted label. The output of the previous convolutional layer is used by the subsequent convolutional layer, which receives the input image from the previous layer and constructs an output feature map.

But, in a DenseNet architecture, all layers are densely connected. That means an inter-layer connection exists between each layer. Moreover, L connections exist between L levels, one between each layer and the layer below it. So, there are $L(L+1)/2$ direct links in the network. The feature maps of all layers before it are utilized as inputs for each layer, and its own feature maps are used as inputs into all levels after it [5]. The DenseNet architecture's dense connectivity can be represented as

$$x(l) = H(l)([x(0), x(0), \ldots, x(l-1)]) \tag{1}$$

SWIN Transformer SWIN transformer is stated as shifted windows transformer. This is basically a hierarchical transformer that is computed with shifted windows. To address the challenges of differences between two domains, such as large variations in the scale of visual entities and the high resolution of pixels in images compared to words in the text, this hierarchical transformer or SWIN transformer was proposed.

In SWIN transformer architecture, first of all, it splits an RGB input image into non-overlapping patches using modules like vision transformer (ViT), and each of the split patches are called "token" which features are set as a concatenation of raw RGB pixel values.

3.4 Brief Work Steps

This study seeks to aggregate locally trained models by retrieving them from local servers. The centrally trained model will be sent to all nearby hospitals following the implementation of federated learning. Most hospitals do not want to share their patient data for privacy issues. So, in our research methodology, we do not take each hospital's datasets, for those hospitals who are willing to connect in this system, we give them a model which is already trained with some test datasets, that model is called a global model. This global model will send to each connected hospital and give them access to fit and train their dataset in that model to contribute to the improvement of accuracy, that model sent to each hospital is called a local model. After successfully fitting and retraining local model, it will be sent to a central server. The server will take the top 80% models based on their accuracy and test whether the model is most accurate than the previous global model if the global model will be overridden by the most accurate local model. The CNN and SWIN transformation algorithms are the foundation of the model. After using the hybrid ensemble CNN model with the VGG19, InceptionV3, and DenseNet algorithms, SWIN transformation was included, and the primary model was developed. Figure. 4 is the whole proposed methodology of our research. First of all X-ray data is being

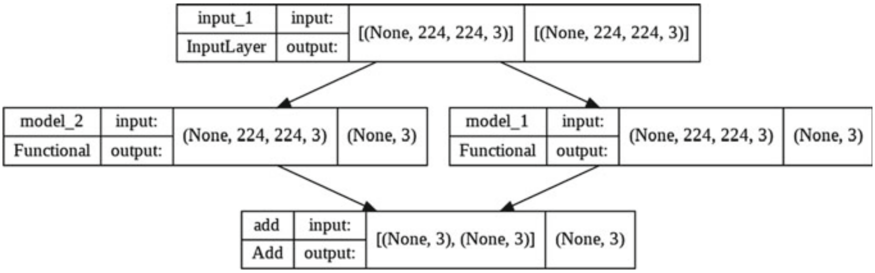


Fig. 3 Transfer and transformer fusion model

collected to make a global model. The dataset is being preprocessed for training in our predictive model, there we have to remove some corrupted image data also and augmentation is being applied.

3.5 Transfer and Transformer Fusion Model

The main part of our model we used hybrid model V1 of VGG19, InceptionV3, and DenseNet201. All the data is being trained separately in VGG19, InceptionV3, and DenseNet201. After the ensemble process, the accuracy will increase and make the model more reliable. Furthermore, adding SWIN transformer to the main fusion model based on the transfer learning model and transformer model will have the leverage to combine transfer learning and transformer-based learning that will ensure novel factors for the model (Fig. 3).

3.6 Federated Learning Centralized Server

The main hybrid trained model will be our initial global model of FL integrated central server. Then, the server will send the global to each hospital's local devices. After that, the hospitals will have their local model on their local devices and continue to work with that. If any of the hospitals have enough datasets to train or fit again in the local model, they can fit into that and make a request to the central server to update that model with the global model. The central server will check and take the top 80% model with better accuracy and update the global model with a local model whose model's predictive performance is better than the previous global and other local models. This work will be done in a loop whenever any hospital will request to update the global model.

4 Implementation and Result Analysis

Disease prediction is one of the most sophisticated examples of advanced computational ability. Now, it is possible to analyze and detect diseases based on CT-image and X-ray images. Thanks to the advancement of artificial intelligence. There are several AI-based models that can do the job of prediction. We have used some cutting-edge technology to predict the difference between COVID-19, pneumonia, and normal X-ray images. We have used VGG19, InceptionV3, DenseNet 201, and SWIN transformer to create our model that can provide reliability to medical practitioners. We have tried to come up with the best approach that can help the medical sector from our computer science field. Moreover, we compared different outcomes that helped us in the way of making the hybrid model more advanced compared to the existing disease detection model with much reliability.

4.1 Implementation

VGG19 VGG19 is the latest pre-trained model from VGGNet architecture. It is the updated version of VGG16. The size of each layer is now 47 which was 41 before in VGG16. Also, it has variants of filter sizes 64, 128, 256, and 512. In our VGG19 model, we have added 3 batch normalization layers along with two dense layer sizes of 128 and 64. Moreover, we also made the middle layers trainable false so that we can avoid overfitting issues. Also, we added a dropout layer with a value of 0.5 to make sure that our model is safe from the overfitting problem. In addition, we have used image sizes (224, 224,3) for our overall processing steps. We have been careful during the choice of size considering our processing unit capability and required time to complete the training without any issues. In addition, during training time, we have calculated steps per epoch and the number of epochs based on the available number of images that also keep our model safe from overtraining. Moreover, with all these careful steps, we have found 94.4 validation accuracy during our training time.

DenseNet201 Now, we start to work with another important convolutional neural network model known DenseNet 201. It is one of the latest neural network architectures available that helped us to make our model even more reliable consisting of 201 layers. We kept the image size (224, 224, 3) during our model training procedure. Along with this, to avoid overtraining, we have made the internal layer trainable to false and added a dropout layer of value 0.5. Next, we used a sequential model as our backbone architecture to pass the DenseNet 201 layers and make our custom model. We have added the global average pooling 2D layer and a dense layer with a size of 1024 to complete the process of designing our custom DenseNet 201 model for improved performance considering the fundamental DenseNet 201 model. In addition, during training time, we have calculated steps per epoch and the number of epochs based on the available number of images and batch size count to keep

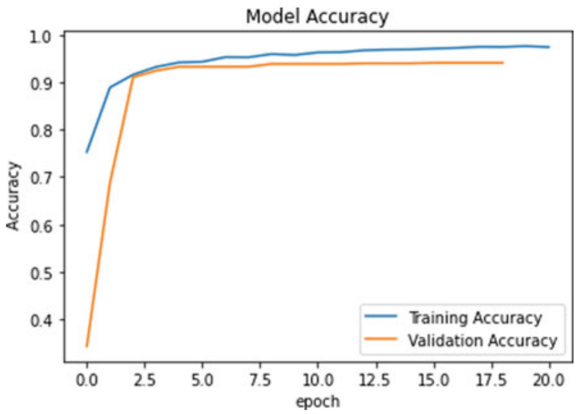


Fig. 4 VGG19 training accuracy and validation categorical accuracy

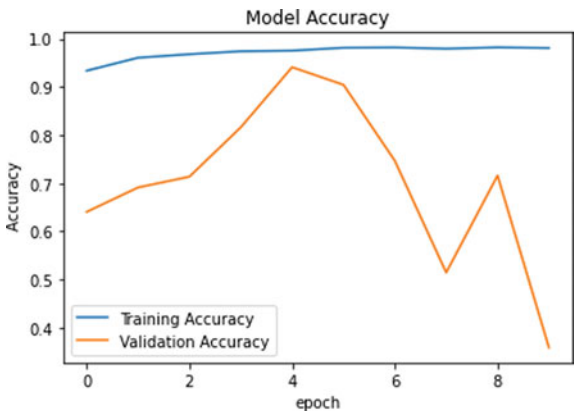


Fig. 5 DenseNet201 training accuracy and validation categorical accuracy

our model safe from overtraining. With all these careful steps, we have found 94.1 categorical validation accuracy during our training time (Fig. 5).

InceptionV3 InceptionV3 is the third edition of Google's inception convolutional neural network. We have used the latest pre-trained model for our disease detection system. InceptionV3 is a parallel processing architecture. The default input image size is (299, 299, 3). However, we used (224, 224, 3) like our previously used model VGG19, DenseNet201. We have used sequential model as our backbone architecture during the implementation of the InceptionV3 model. We have added the global average pooling 2D layer and a dense layer size of 1024 during the development of our custom InceptionV3 model. We have added a dropout layer with a value of 0.5 to avoid over-fitting problems. Furthermore, we have calculated steps per epoch which is 200, and the number of epochs is 30 based on the image count of more than six thousand and batch size of 25. We have fine-tuned the training structure keeping in mind our hardware limitation and avoiding overtraining. With all these careful steps, we have found 94.5% of validation accuracy during our training time.

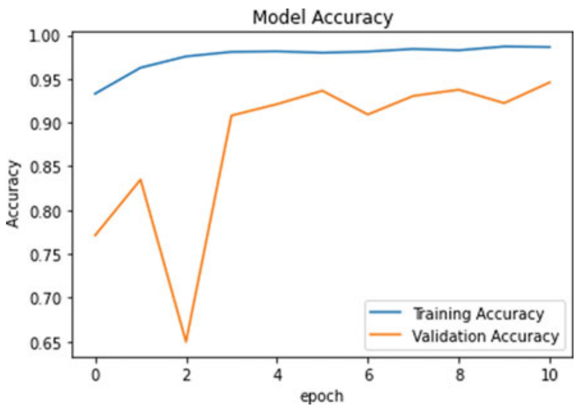


Fig. 6 InceptionV3 training accuracy and validation categorical accuracy

SWIN Transformer SWIN transformer is the CNN architecture with the origin branch of the transformer-based learning approach. It is one of the most prominent architectures developed by Microsoft. The full form of SWIN is shifted window. In this process, we can reach the pixel-level image detail of an image. This transformer learning technique divided the image into different patches before sending it for training. Like the existing CNN model, the SWIN transformer is a large encoder–decoder block that processes the input data. SWIN transformer is the general purpose backbone of computer vision. The shifting window technique brings astonishing efficiency by limiting self-attention computation to non-overlapping local windows while also allowing the cross-window connection. We have used a patch size of (2,2) and a number of attention heads of 8. Moreover, we have used a window size of 7 with a shift size of 1. In our training structure, we have maintained the image dimension is (224, 224, 3). Furthermore, we have calculated steps per epoch that is 200, and the number of epochs is 10 based on the image count of more than six thousand and batch size of 25. We have fine-tuned the training structure keeping in mind our hardware limitation and avoiding overtraining. With all these careful steps, we have found 82.5% of validation accuracy during our training time (Fig. 6).

Fusion Model We have individually trained and tested our discussed models. Now in order to create the hybrid model. The hybrid model is the combination of transfer learning-based models (VGG19, InceptionV3, DenseNet 201) and the transformer model (SWIN Transformer). We have developed this hybrid model. We have used the ensemble technique to combine the entire model and build a single unique model that will work as the backbone of our disease detection system.

We are hopeful that this system that we have worked with COVID-19, pneumonia, and normal X-ray image will also provide significant outcomes if this model is going to use for any other detection system development. Furthermore, we have used the previous training configurations in order to maintain the proper collaboration of the different models and ensure the best throughput of the hybrid model. We have used (224, 224, 3) image size with a dropout layer value of 0.5. Next, we have maintained

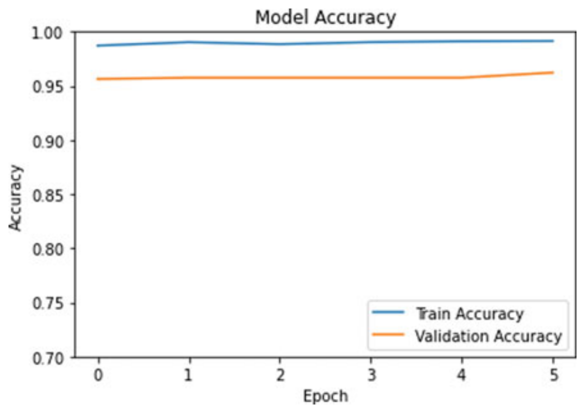


Fig. 7 Training accuracy and validation accuracy (fusion model)

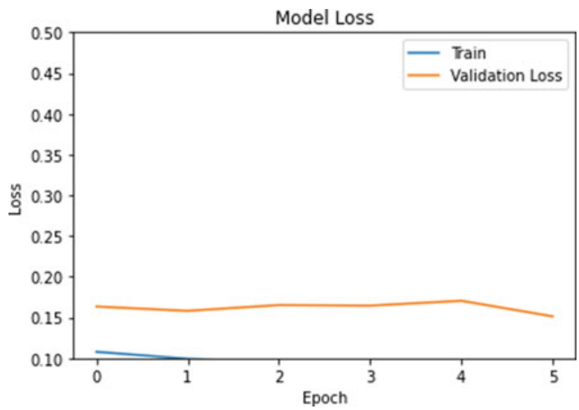


Fig. 8 Training loss and validation loss

a proper training structure consisting of the number of epochs, steps per epoch, and batch size. With all these careful steps, we have found 97.0% of validation accuracy during our training time by combining techniques that sum the weight of all four models. Next, we found 94.0% of validation accuracy by combining techniques that average the weight of all four models.

In Fig. 7, we have shown our trained dataset using fusion model. We have found training accuracy 99% approximately and categorical validation accuracy of 96% approximately. This indicates a slight over-fitting problem. However, we are working to improve this over-fitting problem in our upcoming research outcome.

In Fig. 8, We have found a training loss of approximately less than 0.10% and a categorical validation loss of less than 0.14% approximately using fusion model. This low loss indicates that our model is overfitting despite using middle-layer trainable false and dropout layer values of 0.5.

4.2 Performance Analysis

In Fig. 9, we have shown our findings when we deployed our newly developed algorithm in a federated environment. We did not find an impressive outcome as we were facing hardware resource limitations. In a single run, the federated environment consumed 35GB of RAM.

In Fig. 10, the VGG19's ROC-AUC curve shows a good area under the curve than the NoSkill straight line diagonal. Moreover, in InceptionV3, ROC-AUC curve shows less area under the curve than the VGG19 model. So the InceptionV3 is less accurate than the VGG19 model. Furthermore, DenseNet 201 the value of AUC is also well, but initially, the area under the curve got shrinks but the ultimate area is good.

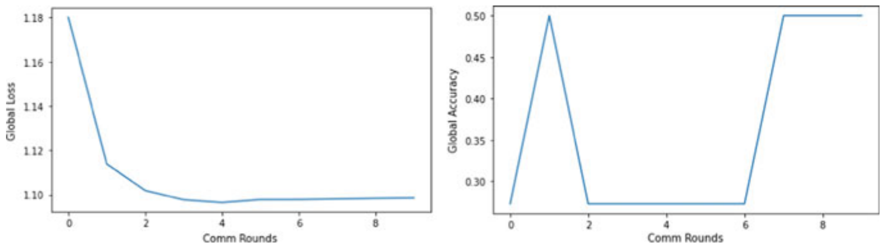


Fig. 9 Federated learning-based outcome

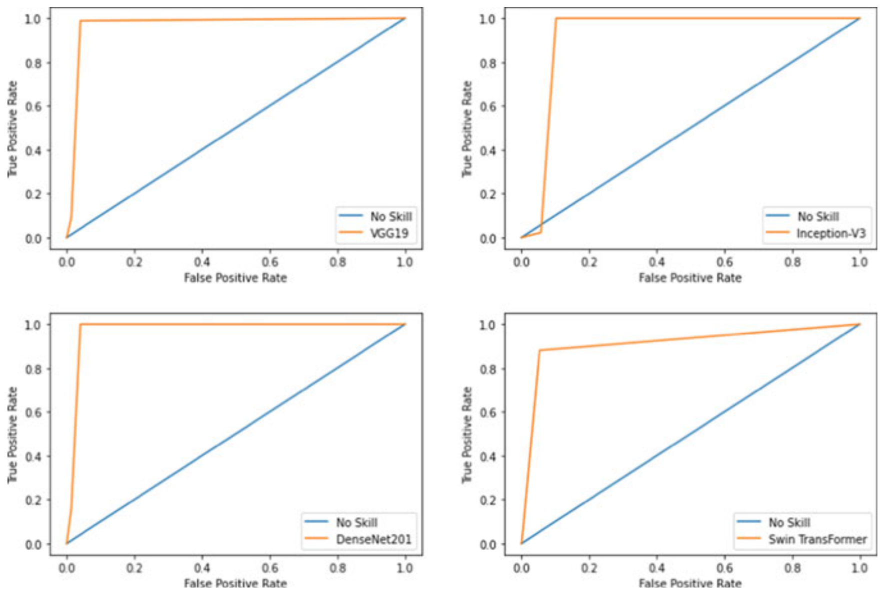


Fig. 10 VGG19, InceptionV3, and DenseNet201 model's AUC-ROC outcomes

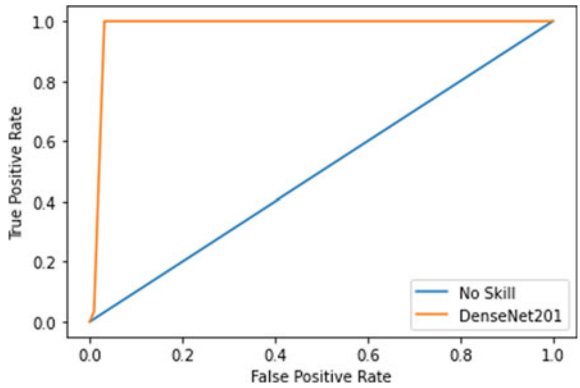


Fig. 11 Fusion model's AUC-ROC

But in Fig. 11, this is the fusion model's AUC-ROC curve; here after combining all the models, we get a good AUC-ROC curve with well AUC value than the other models individually.

Comparative Analysis In Fig. 12, our comparative confusion matrix analysis has been shown between VGG19, InceptionV3, and DenseNet201. All individual models are predicting mostly true-positive. But in Fig. 10, our fusion model's true-positive performance is better than each individual model (Table 1).

We have observed different performance criteria. Among them, we have found the confusion matrix-based test outcome that gives us an idea about the performance comparison (Fig. 13).

5 Conclusion

In this research, we have provided brief explanations on **A hybrid FL-enabled ensemble approach for lung disease diagnosis leveraging a combination of SWIN transformer and CNN.** We have used a combined model of transfer learning and transformer learning known as shifted window transformer to make our model reliable. In future, we want to work on the concept drift of federated learning to address the limitation of federated learning. In addition, we want to improve our algorithm to make it more efficient and reliable during the analysis process of patient data. We want to improve our individual algorithm's efficiency and use better hardware to train our models that will be efficient and reliable. We also want to analyze our program's time and space complexity. In addition, we want to analyze our dataset from a statistical aspect as well. Lastly, we want to use our developed system in different disease detection processes to contribute to the medical sector. However, technology will always work as assistance to medical treatment but will be limited because of the variance of diseases and treatment processes.

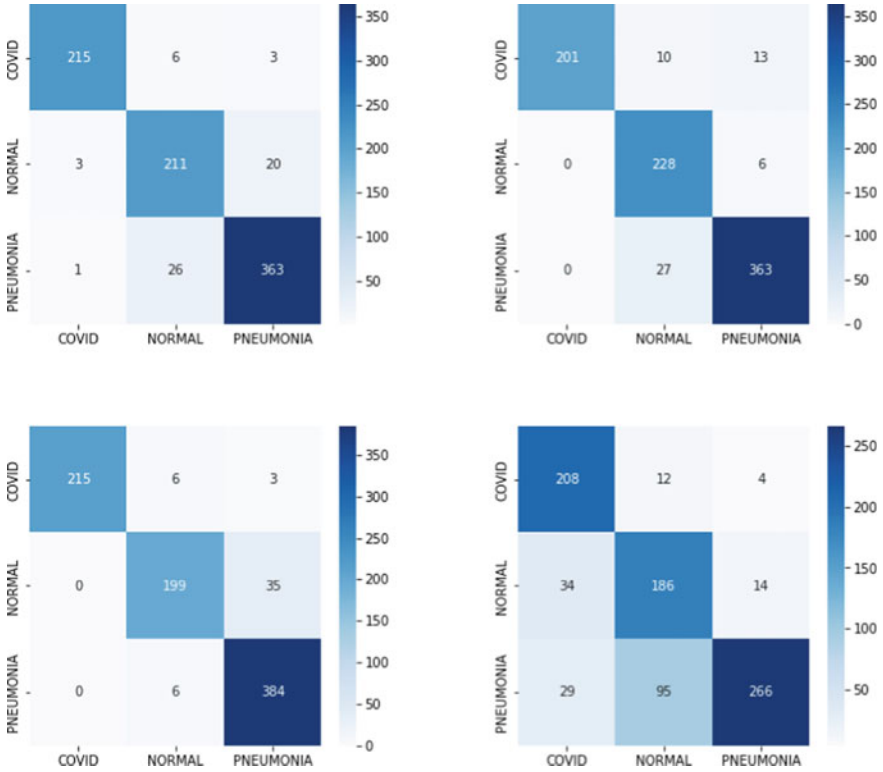


Fig. 12 **a** VGG19 (top-left), **b** InceptionV3 (top-right), **c** DenseNet 201 (bottom-left), and **d** SWIN transformer (bottom-right) model's confusion matrix outcomes

Fig. 13 Fusion model's confusion matrix outcomes

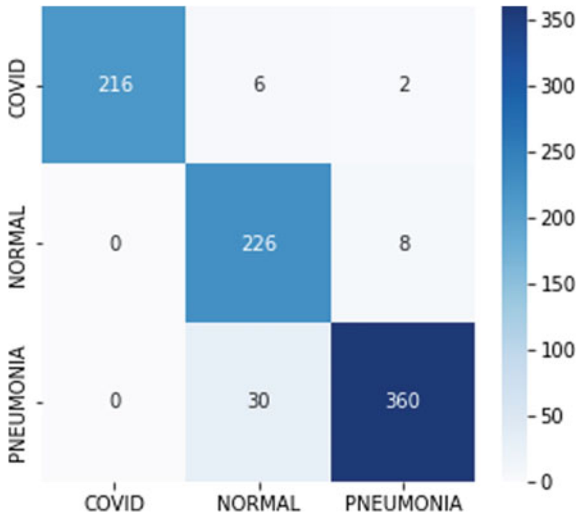


Table 1 Model comparison table

Model comparison			
Classifier	Training time (s) (approx.)	Testing time (s) (approx.)	Accuracy (%)
VGG19	14440	4	94.4
InceptionV3	15200	2	94.5
DenseNet 201	18120	2	94.1
SWIN transformer	25650	4	82.5
Fusion model (sum)	24122	3	96.24
Fusion model (average)	21600	2	94

References

1. Kassania SH, Kassanib PH, Wesolowskic MJ, Schneidera KA, Detersa R (2021) Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach. Biocybern Biomed Eng 41(3):867–879
2. Hemdan EED, Shouman MA, Karar ME (2020) COVIDX-Net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images. arXiv:2003.11055
3. Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ (2020) Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning. Med Image Anal 65:101794
4. Jiang J, Lin S (2021) COVID-19 detection in chest X-ray images using SWIN-transformer and transformer in transformer. arXiv:2110.08427
5. Gu Y, Piao Z, Yoo SJ (2022) STHarDNet: SWIN transformer with HarDNet for MRI segmentation. Appl Sci 12(1):468
6. Li T, Sahu AK, Talwalkar A, Smith V (2020) Federated learning: challenges, methods, and future directions. IEEE Signal Process Mag 37(3):50–60
7. Corinzia L, Beuret A, Buhmann JM (2019) Variational federated multi-task learning. arXiv:1906.06268
8. Smith V, Chiang CK, Sanjabi M, Talwalkar AS (2017) Federated multi-task learning. In: Advances in neural information processing systems, vol 30
9. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, Ramage D, Segal A, Seth K (2017) Practical secure aggregation for privacy-preserving machine learning. In: Proceedings of the 2017 ACM SIGSAC conference on computer and communications security, pp 1175–1191
10. Hemdan EE-D, Shouman MA, Karar ME (2020) COVID-Net: a frame-work of deep learning classifiers to diagnose COVID-19 in X-ray images. arXiv:2003.11055

A Sensor System for Stair Recognition in Active Stair-Climbing Aid: Preliminary Research

Ga-Young Kim, Won-Young Lee, Dae-We Kim, Joo-Hyung Lee, Se-Hoon Park, and Su-Hong Eom

Abstract Since the 2000s, the use of wheelchairs has been increasing to guarantee their right to move as the number of people with walking disabilities continues to increase. However, movement of wheelchairs on the stairs is constrained because they are structurally characterized by two large wheels and two small auxiliary wheels. Therefore, this study proposes an algorithm for estimating the alignment state of stair-climbing aid and the height of the stairs. Also, it proposes a sensor system for estimating the entry section of a flatland. As a result of the experiment, based on the proposed sensor module, the alignment state of the stair-climbing aid is confirmed, and the height of the stairs is estimated by detecting the edge of the stairs. Also, the entry position of a flatland at the end of the stairs is detected.

Keywords Wheelchair stair-climbing · Stair entry angle · Stair-climbing aid

---

G.-Y. Kim (✉) · W.-Y. Lee · D.-W. Kim · S.-H. Eom
Department of Electronic Engineering, Tech University of Korea, Siheung, Republic of Korea
e-mail: rk007954@tukorea.ac.kr

W.-Y. Lee
e-mail: lbborie@tukorea.ac.kr

D.-W. Kim
e-mail: daewek96@tukorea.ac.kr

S.-H. Eom
e-mail: sheom@tukorea.ac.kr

J.-H. Lee
Department of Computer Engineering, Tech University of Korea, Siheung, Republic of Korea
e-mail: patrickjumba@tukorea.ac.kr

S.-H. Park
Korea Orthopedics and Rehabilitation Engineering Center, Incheon, Republic of Korea
e-mail: mbb1020@kcomwel.or.kr

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_78

1 Introduction

As the global aging population and the number of people with walking disabilities continue to increase since the 2000s, aids to guarantee their right to move have become a social topic [1]. People who have difficulty walking in their daily lives in the United States and Europe, Korea, are called mobility handicapped persons. The number of wheelchair users, a representative aid that supports them, is constantly increasing. With this increase in demand, the global wheelchair market is expected to reach $12.2 billion by 2030, according to Grand View Research, Inc. in 2022.

Due to the continuous expansion of the wheelchair market, technology that values convenience and stability of wheelchair is currently required beyond popularization. Therefore, response technologies and autonomous driving technologies for convenience and safety of operation have made rapid progress since the 2000s [2]. However, these are not technological advances to enhance the freedom of wheelchair movement.

Wheelchairs are structurally characterized by two large wheels and two small auxiliary wheels. Therefore, movement on the stairs is constrained. Movement on the stairs requires the use of separate elevator or lift. In the case of a building without an elevator or lift, there is a difficulty that the guardian must lift the wheelchair in order to move on the stairs [3]. In addition, additional costs are required because elevator or lift must be installed on each route, and it may be difficult to install due to structural problems [4].

Several products have been developed to improve this problem, such as iBot, a convertible four-wheel inverted pendulum structure wheelchair, Scalevo and Topchair-S, an electric wheelchair with track wheels, and are currently on the market. However, the high price and all-in-one electric wheelchairs and track wheel modules did not lead to popularization. Therefore, there is a growing need to develop a stair-climbing aid that can move on the stairs just by attaching a device to a manual wheelchair that is easier to move than an electric wheelchair.

Currently, the representative stair-climbing aid operated by the guardian is an infinite-track product such as the Liftkar PTR manufactured by SANO. This platform has excellent contact performance of curved surfaces due to the projection of the track, so the ground area is evenly distributed, enabling safe driving of obstacles or stairs [5, 6]. However, the guardian is generally a family of wheelchair users. If the guardian is old, it is not easy to use the product currently on the market. To use it, the platform must be properly positioned to climb the stairs to ensure straightness, and there is a difficulty in ensuring that the platform is adequately supported to prevent it from falling at the end of the stairs.

Since existing platforms have limitations, it is necessary to adopt a convertible double-wheel caterpillar structure or landing gear. Prior to that, this study conducts a preliminary study on the sensor system that recognizes the situation of stair-climbing. Therefore, this study proposes a ToF sensor-based stairs recognition method and an algorithm for estimating the alignment state of stair-climbing aid. Also, it proposes a sensor system for estimating the entry section of a flatland.

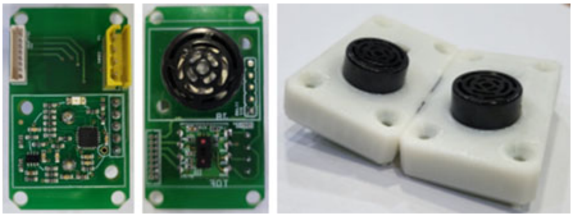


Fig. 1 Proposed sensor fusion module

2 Research Method

2.1 Sensor System for Estimating the Position of the Stairs by the Stair-Climbing Aid

In this section, a sensor module for estimating the position of the stairs is proposed. The two ToF sensors are combined to estimate the position of the stairs. They are ultrasonic sensors and multi-zone ranging ToF sensors. Based on this, an algorithm for estimating the alignment state of the stair-climbing aid in front of the stairs and the height of the stairs are proposed.

2.1.1 Combined Sensor Module using Ultrasonic Sensors and Multi-zone Ranging ToF Sensors

It is difficult to estimate the misalignment state of an object using a single ultrasonic sensor due to the limitation of ultrasonic radiation angle and sound wave diffraction. On the other hand, the multi-zone ranging ToF sensor has a maximum of 8×8 multi-zone based on 940 nm invisible IR light. And it is possible to estimate the misalignment state using a single sensor compared to an ultrasonic sensor, which uses several sensors. The ToF sensor is VL53L5CX, STMicroelectronics [7]. However, as the distance from the measurement target increases, the area of the object recognized in the multi-zone may lead to unexpected results. Therefore, as shown in Fig. 1, this study proposes one sensor module in order to use the advantages of ultrasonic sensor and multi-zone ranging ToF sensor.

2.1.2 Algorithm for Estimating the Alignment State of the Stair-Climbing Aid and the Height of the Stairs

The algorithm for estimating the alignment state in front of the stairs of the stair-climbing aid is designed as shown in Fig. 2. If the distance between the platform and the stairs measured by the ultrasonic sensor is a certain threshold value, measurement value of the ToF sensor is obtained and the values of each zone are compared. It allows the detection area to be measured in a 4×4 or 8×8 matrix. However, the ToF sensor cannot be used at a long distance. Because the IR cannot be reflected depending on the type of object or ambient temperature. If the deviation between each zone is within 5–10%, it is determined that the platform is properly aligned, and the height of the stairs is estimated to start climbing the stairs.

As for the height of the stairs, the edge of the stairs is estimated by checking the ToF sensor value every 10 minutes while raising the landing gear or double-wheel caterpillar of the platform. Once the edge is checked, the change in position where the sensor module attached to the platform moved from the ground is measured based on the encoder sensor or other sensors mounted on the platform. And it is calculated the entry angle that is easy to climb the stairs.

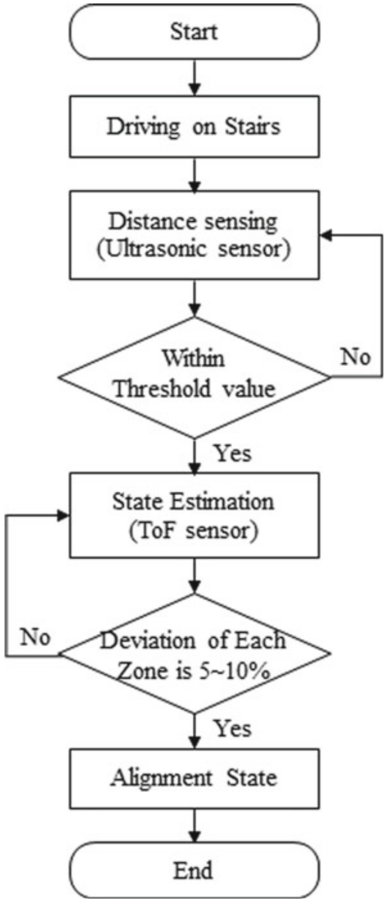


Fig. 2 The algorithm for estimating the alignment state in front of the stairs of the stair-climbing

2.2 Sensor System for Estimating the Entry Section of a Flatland of the Stair-Climbing Aid

In this section, a sensor system for detecting the safe entry section of a flatland at the end of the stairs is proposed. In order to estimate entry section of a flatland, the pressure sensing unit is designed based on a pressure sensor that is strong in the external environment. And the entry section of a flatland detection algorithm is designed using it.

2.2.1 Pressure Sensing Sensor Module

To detect the entry section of the flatland at the end of the stairs while climbing the stairs, a pressure detection unit including a pressure sensor such as a switch on the belt support is designed, as shown in Fig. 3. The pressure sensing unit is located on the bottom surface of the belt support. And it detects the pressure applied to the belt support by the edge of the stairs during the stair-climbing. The pressure sensing unit is located at the front and rear of the belt support, and can detect up to two stair corners during stair-climbing. The pressure sensing unit is designed to be located at the edge of the stairs with one step difference based on the range of stairs presented in the Ordinance No. 548 of the Ministry of Land, Infrastructure, and Transport of Korea. In addition, in order to apply to various types of stairs, the sensing area of the pressure sensing unit is designed to be wider than the sensing area of the pressure sensor in order for detecting a wide range of pressure. The pressure sensing unit is manufactured by attaching a spring elastic body between the belt support and the pressure sensor to determine whether there is pressure or not.

2.2.2 Algorithm for Estimating the Entry Section of a Flatland

The algorithm for estimating the entry section of a flatland is designed as shown in Fig. 4. The pressure sensor of the pressure sensing unit detects pressure during the stair-climbing and determines whether it moves on the edge of the stairs. If pressure is applied to the front pressure sensing unit after the pressure is detected in the rear pressure sensing unit, it is a climbing situation on the stairs. If the pressure is not detected in the front pressure sensing unit after the pressure is detected in the rear pressure sensing unit, it is determined as the entry section of the flatland. If the location of the stair-climbing aid is estimated to be the flatland entry section through the pressure sensing, the landing gear or the double-wheel caterpillar is controlled to safely enter the flatland. It is expected that the pressure sensing unit and the flatland entry section detection algorithm will improve the flatland entry section estimation problem caused by external environments, thereby ensuring the safety of the stair-climbing aid operator.

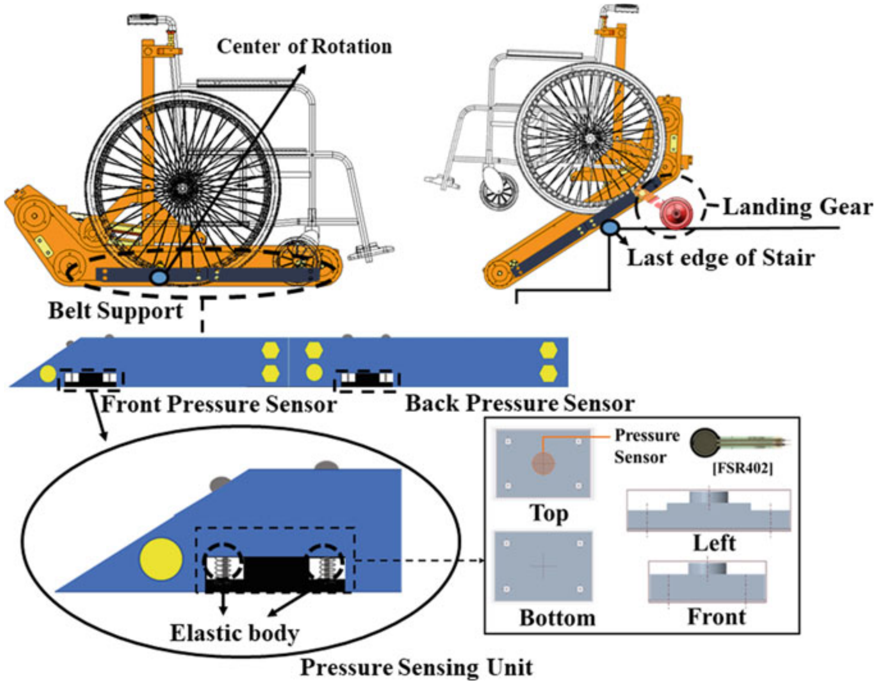


Fig. 3 Structure of the pressure detection sensor system on the stair-climbing aid

3 Experiments

3.1 *Experiments for Estimating the Alignment State of the Stair-Climbing Aid and the Height of the Stairs*

The results of estimating the alignment state of the stair-climbing aid are shown in Figs. 5 and 6. The platform is 30 cm away from the stairs. In the alignment state of the platform, the measurement result shows that the deviation of each zone is within 5–10%. However, the measurement result shows that the deviation of each zone is large while it is 20° twisted from the front of the stairs. Therefore, it is confirmed that the alignment state of the platform may be estimated through a sensor module combined with ultrasonic sensors and multi-zone sensing ToF sensors.

The results of detecting the edge of the stairs to estimate the height of the stairs are shown in Fig. 7. Likewise, the platform is 30 cm away from the stairs. It can be seen that the value measured at the edge of the stairs is the smallest. Because the distance between the platform and the edge of the stairs is the shortest. In addition, it can be seen that the value of each zone is measured gradually from the edge of the stairs. Therefore, it is confirmed that the edge of the stairs can be detected by the multi-zone sensing ToF sensors.

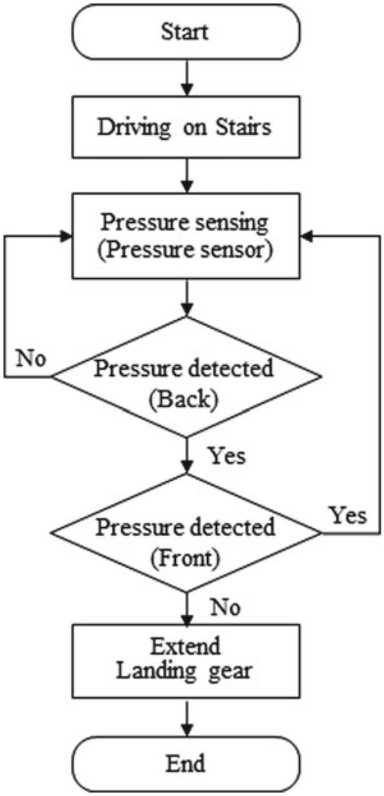


Fig. 4 Algorithm for detecting the change section of the stair flatland

Fig. 5 Estimation of the alignment state of the stair-climbing aid

570	447	525	546	545	530	487	X
355	339	340	334	334	346	363	X
327	319	322	319	317	316	320	315
320	317	318	319	315	316	316	313
319	317	315	318	315	315	311	309
318	314	314	315	311	312	311	310
310	310	308	309	311	308	306	309
251	258	276	279	280	280	278	274

3.2 Experiments for Estimating the Entry Position of a Flatland

The result of the front and rear pressure sensing unit during the stair-climbing is shown in Fig. 8. It can be seen that pressure was detected only in the front pressure sensing unit when entering the stairs. It is confirmed that the wheelchair is

Fig. 6 Estimation of the misalignment state of the stair-climbing aid

X	876	960	592	534	544	532	453
663	597	532	X	346	332	334	352
X	X	X	X	322	315	303	294
361	350	341	331	322	312	302	290
361	347	340	333	319	313	301	290
355	348	338	327	320	307	299	289
340	335	332	320	312	303	297	287
243	254	269	290	283	281	276	270

Fig. 7 Edge detection of the stairs

484	479	476	480	485	478	482	465
493	487	485	483	483	482	495	495
452	401	345	341	335	359	373	394
277	273	268	264	264	268	271	268
287	282	273	269	269	275	278	275
278	278	278	272	276	275	276	274
285	283	283	284	282	280	283	281
296	290	290	292	292	288	291	293

climbing the stairs by detecting the pressure at the rear pressure sensing unit after detecting the pressure at the front pressure sensing unit. When moving the flatland entry section, however, the pressure was not detected at the front pressure sensing unit after detecting the pressure at the rear pressure sensing unit. Then, it was judged by the edge of the last step and was confirmed that the landing gear was controlled to safely enter the flatland.

Figure 9 shows the acceleration data applied to the user before and after applying the flatland entry section detection method. The maximum acceleration was 1.15[g] when entering the flatland at the user boarding area before applying the method, and a maximum acceleration of 0.55[g] was detected after applying the method. In this way, the impact amount decreased by 47.8% as the maximum acceleration applied to the user was reduced compared to before the method was applied, and the safety was increased to confirm the validity of the flatland detection method.

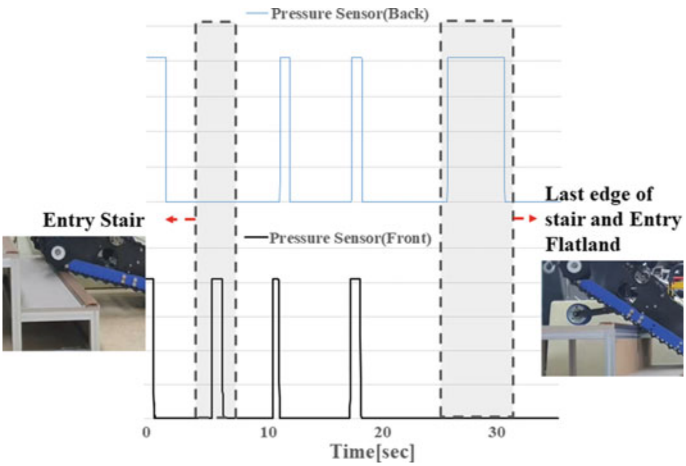


Fig. 8 Estimation of the flatland entry position—Data of the front and rear pressure sensing units

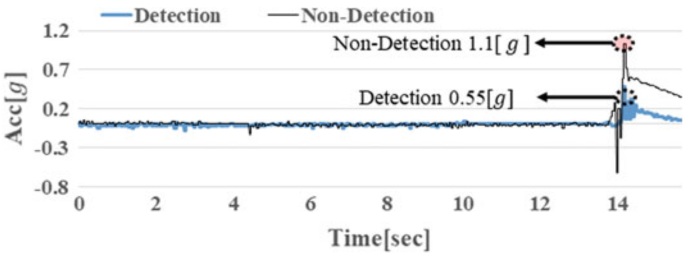


Fig. 9 Estimation of the flatland entry position—Acceleration before and after applying the flatland entry section detection method

4 Conclusions

This study aimed to develop a stair-climbing aid to promote user convenience for moving passive or light wheelchairs in a stair. To this end, sensor modules and algorithms for estimating the alignment state and height of the stairs for the stair-climbing were designed. Also, a sensor system was proposed to estimate the entry position of a flatland. In addition, the validity of the proposed method was confirmed through a quantitative verification process, and the contents of this study are expected to be used for the activation of the existing wheelchair stair-climbing aid.

Acknowledgements "This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare, Republic of Korea (HJ20C0058)."

"This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2022-2018-0-01426) supervised by the IITP (Institute for Information & Communications Technology Planning and Evaluation)"

"This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2022- 2020-0-01741-003) supervised by the IITP (Institute for Information & communications Technology Planning and Evaluation)."

References

1. Endo D, Watanabe A, Nagatani K (2016) Stair climbing control of 4 degrees of freedom tracked vehicle based on internal sen sors. In: 2016 IEEE international symposium, pp 112–117
2. Lee L-K, Se-Young O (2015) Development of smart wheelchair system and navigation technology for stable driving performance in indoor-outdoor. Environ J Inst Electron Inf Eng 52(7):1377–1385
3. Chang HH, Lee WY (2019) A study on the recognition of stairway steps in wheelchair movement assistants. Inf Control Sympos pp 77–90
4. A Survey on the Safety Status of Facilities for the Handicapped in Subway. Korea Consumer Agency, October 2018
5. Cho HS, Ryu JC (2011) Development of driving simulation model for stair climbing wheelchair. In: Korean society for precision engineering, pp 1401–1402
6. Guillon LB, Fermanian C, Pouillot S, Boyer F et al (2008) Evaluation of a stair-climbing power wheelchair in 25 people with tetraplegia. ArchPhys Med Rehabil, pp 1958–1964
7. Niculescu V, Muller H, Ostovar I, Polonelli T, Magno M, Benini L (2022) Towards a multi-pixel time-of-flight indoor navigation system for nano-drone applications. In: 2022 IEEE international instrumentation and measurement technology conference

Decentralised Renewable Electricity Certificates Using Smart Meters and Blockchain

Yuki Sato, Szilard Zsolt Fazekas, and Akihiro Yamamura

Abstract Recently, the use of renewable energy has been promoted to decarbonise the world. Renewable energy certificates are issued to prove the authenticity of source of renewable energy. However, it is not easy to deal with this system because of security and privacy issues caused in the process amongst stakeholders. We analyse plausible security and privacy issues caused in related ICT systems and then propose a renewable electricity certificate to be issued in a decentralised manner using smart meters and a blockchain.

Keywords Blockchain · Smart meters · Renewable energy certificates · Security · Privacy

1 Introduction

Climate change and global warming attract much attention because they severely affect the quality of our life. The current rise in global average temperature is more rapid than the previous changes and is believed to be primarily caused by human activity. Burning fossil fuels (coal, oil, and natural gas [1]) increase emission of greenhouse gases, notably carbon dioxide. Climate change can be mitigated by reducing greenhouse gas emissions; in particular, reducing use of fossil fuels is desired worldwide as discussed in COP27 [2]. Fossil fuels accounted for 80% of the world's energy, and the remaining share is comprised of nuclear power and renewable energies including hydropower, bioenergy, wind and solar power, and geothermal energy. Renewable energy plays an important role in slowing down carbon dioxide emissions.

Y. Sato (✉) · S. Z. Fazekas · A. Yamamura
Department of Mathematical Science and Electrical-Electronic-Computer Engineering, Akita University, Akita, Japan
e-mail: yukisato.akita@gmail.com

S. Z. Fazekas
e-mail: szilard.fazekas@ie.akita-u.ac.jp

A. Yamamura
e-mail: yamamura@ie.akita-u.ac.jp

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_79

In several countries, a carbon tax is levied on the carbon dioxide emissions caused by producing goods and services. It is believed that a carbon tax reduces carbon dioxide emissions by increasing prices of the fossil fuels that emit them when burned. A crucial issue is how to make visible the social costs of carbon dioxide emissions caused by human activity, otherwise we do not know how to estimate emission of greenhouse gases and compute the tax amount accordingly. We have not established a fair and transparent process to estimate carbon dioxide emissions through human activities so far. One of the challenges to this problem is a renewable energy certificate which is a means to prove the source of renewable energy. It has been already implemented in several countries although no uniform regulations are set amongst countries. For example, renewable energy certificate (US), green certificate (Europe), and Renewable Energy Certificates Registry and GreenTag (Australia) have been installed. These systems may be suffering from some security flaws, in particular, the big brother problem, that is, digital surveillance by government and private actors and its implications for human rights. This problem appears when the system is constructed based on centralised control over the participants.

Renewable energy certificates can be used, for example, to prove that a product is made exclusively from renewable energy sources to reduce the tax payment and are only issued by each organisation related to an energy source or a country. These systems work only if the organisation does not act in any malicious way. It is not easy for ordinary households to issue a certificate and deal with regulation. Therefore, a third-party organisation issues a certification instead. On the other hand, this causes serious security and privacy threats. Some measures for information security are expected for adequate applications of a renewable energy certificate.

The certificates are issued by a third-party organisation that sells electricity with certificates. The purchasers of electricity with certicicate can use the certificate to prove that electricity is generated from renewable energy sources when they sell their products. However, they may get more profit by cheating (Table 1).

Table 1 Data in renewable energy certificate

Data	
Amount of electricity	Amount of electricity to which the certificate applies
Period of application	Period during which the certificate can be applied
Maker information	Indicates whether the certificate was produced correctly
Generation data	Indicates the distribution channel
Owner information	Expresses delivery of certificates

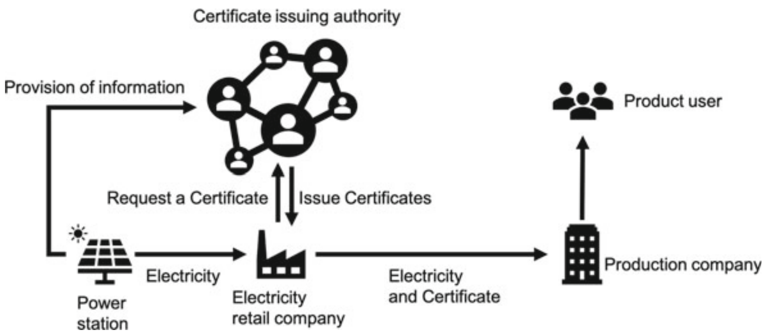


Fig. 1 Proposed model

Table 2 Stakeholders

Stakeholder	Roll
R: Retail company	Buy and sell electricity Produces certificate data
T: Third-party organisation	Authenticate certificates
P: Power plant	Generate and sell electricity
U: Certificate user	Produce and/or sell products with certificates

2 Models and Plausible Threats

In this paper, we concentrate on *certificates to prove authenticity* of renewable electricity generation, not renewable energy certificates in general. Therefore, we restrict our argument to the case when power generators are fixed and power plants accumulate this electricity. An outline of our proposal is shown in Fig. 1. Plausible stakeholders are retail companies, third-party organisations, power plants, and certificate users as summarised in Table 2. Each of the stakeholders has motivation for malicious acts such as misconduct, dishonesty, and cheating for economic reasons. In addition, outsiders may act malicious behaviours for economical reasons, however, we do not touch on the issue in this paper. We shall introduce countermeasures to prevent these malicious acts.

First, we discuss malicious acts of the stakeholders; **R**, **T**, **P**, and **U** stand for a retail company, a third-party organisation, a power plant, and certificate user, respectively. A retail company **R** purchases and sells electricity, and **R** produces a certificate data at the same time. Note that **R** can make more profit by selling renewable energy and so has motivation to falsify certificates; **R** can sell electricity generated from fossil fuels in addition to renewable energy and make a profit by rewriting the amount of electricity generated on the certificate. A certificate user **U** produces and sells products with a certificate guaranteed that it is made exclusively from renewable energy. This helps to exempt carbon taxes. However, the certificate may be falsified, that is, fossil fuel electricity instead of renewable electricity may be used in the manufacturing process. A power plant **P** generates and sells electricity. Note that **P**, like **R**, can fraudulently make profit by forging certificates saying the electricity is generated from renewable sources, when it is not. We remark that electricity generated from renewable energy sources often costs much more than fossil fuels [3]. A third-party organisation **T** authenticates certificates and can forge certificates just to disturb **R**'s business and obtain inside information which is hidden. This is caused by allowing only one trusted party as **T** to deal with issuing certificates. Authentic certificates may be used twice during their validity period unless some checking system is installed. In summary, certificates are expected to possess:

(1) Mechanisms to check authenticity of certificates.
(2) Decentralised issuance and management of certificates.
(3) Verification system to check the amount of electricity.
(4) Method to trace the source of certificates.

In addition, certificates need to be easily issued, regardless of whether they are for large-scale or small-scale power generation.

3 Decentralised Renewable Electricity Certificates

3.1 Fundamental Technologies: Blockchain and Smart Meters

A blockchain is a technology used for Bitcoin proposed in [4]. It is characterised by its decentralised nature, which is achieved by a consensus-building algorithm called proof-of-work. It is also characterised by the inclusion of the hash value of the previous block to increase tamper resistance. Proof-of-work is a consensus algorithm that uses hash calculations and is performed when a block is included in the blockchain. There are various consensus algorithms for a blockchain other than the proof-of-work [5]. See [6] for detailed information on a blockchain. We employ a blockchain since we need some mechanism to decentralise the role of **R** and to differentiate the proposed scheme from the existing schemes suffering from the big brother problem.

Smart meters are attached to distribution power lines and can transmit real-time electricity information from remote locations. They are used to determine electricity prices in real time [7]. This system is based on the fact that electricity information can be communicated in real time and that the installation of a smart meter simplifies the issuance of certificates even for small-scale power generation like feed-in tariffs (FIT) generation (see [8]).

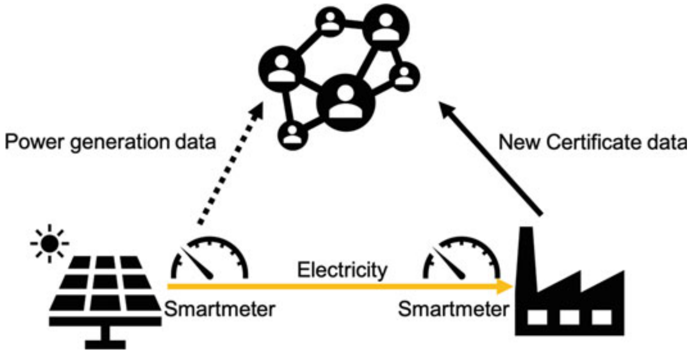


Fig. 2 Process of issuing a certificate

3.2 Proposed Scheme

A power plant **P** sells the electricity generated by renewable resources to **R**, and then **R** creates renewable energy certificate data. Issuance of a certificate is divided into three processes:

(1) Buying and selling electricity, preparation of certificate data.
(2) Verification and broadcast of block containing certificate data.
(3) Receipt and use of certificates.

See Fig. 2.

The electricity purchase and sale process are divided into four processes:

(1) A power plant **P** sells electricity to a retailer $\mathbf{R_1}$.
(2) Transmission of the encrypted data of electricity generated from the smart meter to the data pool.
(3) A retailer $\mathbf{R_1}$ produces certificate data on the purchased electricity and transmits it to representative retailer $\mathbf{R_2}$.
(4) The representative retailer $\mathbf{R_2}$ compiles the certificate data sent to it, creates a block, and sends it to peer-to-peer network.

The system uses smart meters, which prevents cheating in the amount of electricity generated by **P**. The data sent to the peer-to-peer network consisting of **R**s from **P** smart meter is also used to verify certificates. Representative companies are selected for each region and shared within the peer-to-peer network in advance. By selecting the **R** that handles the most electricity in a country or region, the transmission of certificate data can be reduced. The representative creates and verifies the blocks. See Fig. 3.

The data structure of the block is created by certificate data, block creator information, hash value of a previous block, and timestamp. The block stores the certificate data and adds the creator information. The block creator information is used to verify that the correct person is creating the block. The previous block's hash value

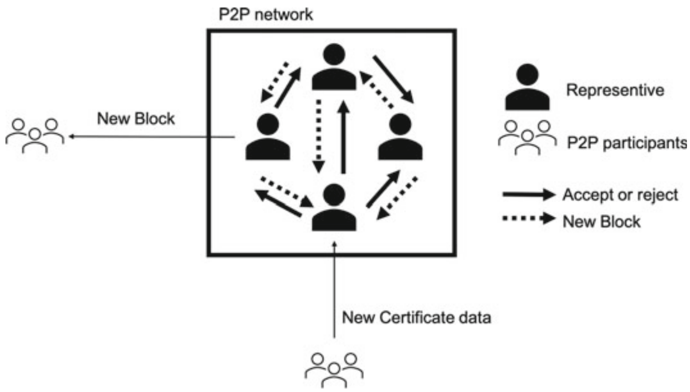


Fig. 3 Process of verifing a certificate

is included to ensure that the value is tamper-resistant. The timestamp can also be added to indicate a correspondence with the date of issue of the certificate. The flow of the block being included in the blockchain is as follows:

(1) A retail company $\mathbf{R_1}$ sends the created certificate data to representative $\mathbf{R_2}$.
(2) $\mathbf{R_2}$ compiles the data sent and creates a block.
(3) The representative verifies the block.
(4) Send the block to the **R** if the verification result is correct and update the blockchain. Verification is done by (1) checking if the block generator is correct and (2) verifying that each value is correct based on the information in the data pool. Decentralisation is achieved by having verification performed by peer-to-peer network representatives.

Of course, the representatives need to be incentivised to perform verification. The verifying **R**s can sell certificates. They can make a profit by including a commission in the price when selling this certificate. Of course, this fee has to be set appropriately. However, there will always be buyers due to the existence of companies like RE100 [9] that have agreed to international initiatives that aim to operate on 100% renewable energy. We believe that the profit is a sufficient incentive to verify correctly.

To prevent double use, all information about the delivery is recorded on the blockchain. The delivery of certificates is agreed on a peer-to-peer network through **R**. A certificate user **U** sends the information about the certificate to be delivered to **R**. Then, **R** creates a new certificate based on that data and sends it to the peer-to-peer network. This is where the change of ownership information takes place. At this point, the digital signature of **R**'s private key expresses agreement to the transfer. The new certificate data is appended with the hash value of the delivered certificate. The addition of this hash value enables tracking. **U** must ensure that it is the correct (Fig. 4).

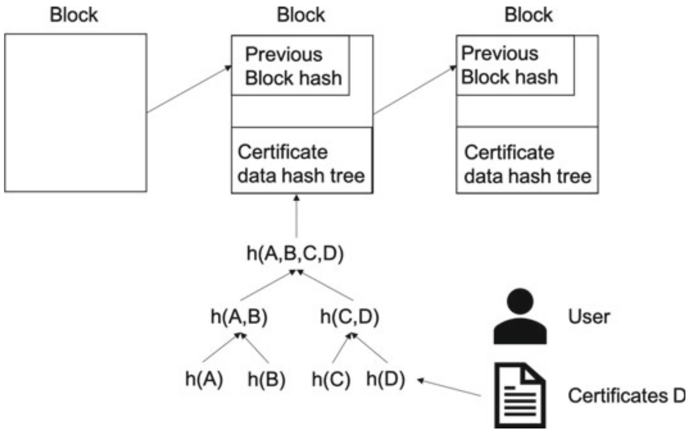


Fig. 4 Verification of authenticity of certificates

Function h represents a hash function such as SHA-256, and A, B, C, and D are the certificate data. **U** can be identified by:

(1) **U** gets $h(C)$ and $h(A, B)$.
(2) **U** calculates $h(C, D)$ and $h(A, B, C, D)$ using $h(A, B)$ and $h(C, D)$.
(3) Verification of the correct certificate by checking whether its figures match those of the Merkle root in the block.

Consider an example of the actual use of a certificate. As an example, let us look at the use of certificates in the purchase and sale of a product created by **U**. At that time, **U** creates a certificate owned by **R** and sends the data to **U**. **U** checks that the certificate is correct if necessary. **U** then creates the product and sells it. When selling a product, **U** compiles the certificate data for the electricity used and requests the peer-to-peer network to deliver the certificate. Here, the certificate is owned by the purchaser of the product. Of course, the purchaser of the product can easily check that the certificate received is correct using the Merkle root.

Consider whether the current model actually prevents cheating as follows:

(1) The rewrite is exposed because the correct certificates are included in the block (if try to rewrite it, need to rewrite all the previous blocks before the one that contains the certificate).
(2) About inadequate certificate issuance: as decentralisation is used, issuance cannot take place. If one tries to make a fraudulent certificate, it cannot be done as it would require the agreement of the other participants.
(3) About double use of certificates: as the delivery of certificates is carried out on the blockchain, they cannot be reused. If double use is attempted, it shows that the certificate has been used previously.

Certificates created under this model can also be diverted for carbon offsetting [10]. As the data relating to the certificates is stored on the blockchain, anyone can check the information. This transparency can also be useful for other applications.

4 Conclusion

In this paper, the issuance and management of certificates were carried out in a decentralised manner using blockchain. The use of smart meters also simplified the issuance of certificates for small-scale power generation like feed-in tariffs (FIT) generation. However, incentives for peer-to-peer network participants need to be discussed. It is important to ensure that such incentives would prevent fraud (e.g. 51% attacks).

References

1. Ellabban O, Abu-Rub H, Blaabjerg F (2014) Renewable energy resources: current status, future prospects and their enabling technology. Renew Sustain Energy Rev 39:748–764
2. The United Nations Climate Change Conference. Available at https://www.cop27.eg/
3. U.S. Energy Information Administration (2022) Cost and performance characteristics of new generating technologies. Annual Energy Outlook
4. Nakamoto S. Bitcoin: a peer-to-peer electronic cash system. Available at https://bitcoin.org/bitcoin.pdf
5. Xiaoqi L, Peng J, Ting C, Xipu L, Qiaoyan W (2020) A survey on the security of blockchain systems. Future Gener Comput Syst 107:841–853
6. Antonopoulos AM (2017) Mastering bitcoin. O'Reilly Media
7. Aswin RC, Aravind E, Ramya SB, Shriram KV (2015) Smart meter based on real time pricing. Proc Technol 21:120–124
8. Couture T, Gagnon Y (2010) An analysis of feed-in tariff remuneration models: implications for renewable energy investment. Energy Policy 38(2):955–965
9. RE100 Climate Group. Available at https://www.there100.org/ja
10. World Resources Institute. Available at https://www.wri.org/research/bottom-line-offsets

The Integration Between Social Media and Customer Relationship Management: The Reliability Analysis

Norizan Anwar, Mohamad Noorman Masrek, Shamila Mohamed Shuhidan, and Yohannes Kurniawan

Abstract Most organization fully utilize social media (SocMed) for their operations in response to Pandemic COVID-19. SocMed, i.e., e-commerce has attracted more users to use and studies found that SocMed manages Customer Relationship and indirectly influences business performance. Customer Relationship Management (CRM) has been used for quite sometimes ago by most organizations. However, studies claimed that organization did not fully utilize the features of CRM. They just keep customer information without doing anything for future strategic planning, competitive advantage, decision-making, and much more. Therefore, this chapter aims to investigate the concept of integration between SocMed and CRM. SPSS version 26 has been used in analyzing the descriptive analysis, i.e., demographic profile and reliability analysis. The results show that all the variables surpass the cutoff value for the pilot test. This pilot study is giving an insight to the researchers in pursuing the data collection.

Keywords Social media · Customer relationship management · Reliability analysis · Integration · SMEs

1 Introduction

Pandemic Covid-19 has triggered all organizations to change their business model. This is to ensure their sustainability and remain competitive in market. The use of social media (SocMed) is seen to have come at the right time. Organizations start to incorporate SocMed in their business activities beside existing systems, i.e.,

N. Anwar (✉) · M. N. Masrek · S. M. Shuhidan
School of Information Science, College of Computing, Informatics and Media, Universiti Teknologi MARA, 40150 Shah Alam, Selangor, Malaysia
e-mail: norizan8027@uitm.edu.my

Y. Kurniawan
Information Systems Department, School of Information Systems, Bina Nusantara University, Jakarta 11480, Indonesia

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_80

Customer Relationship Management (CRM). In year 2020, the statistic shows that the total world population is increased by 1.35%, which is 413 thousand people. The increment led 2.8% for internet users, which is 738 thousand, and 7.7%, which is 2.0 million for active social media users [1]. That is why SocMed becomes top choice by users to purchase goods.

SocMed has the ability to engage their users in their own way, i.e., posts, stories, reels, and live. These features led organizations to use SocMed as a platform for their marketing, reaching out their customers, launching new product(s), product(s) evaluation, and much more. Apart from that, companies are also able to update posts, stories, reels at any hours and event do live anywhere. Organizations can connect with their existing customer more closely and attract potential customer at the same time.

Utilizing SocMed to manage customer relationships can significantly influence the organization's performance due to application's increasing customer engagement and the value created from these engagements [2]. Empirical study has shown that that SMEs' customers use SocMed to generate content, influence other customers through positive reviews, and mobilize others' actions toward the brands or products [3]. According to [4], SocMed increases communication around brands and products, enhances positive as well as negative word-of-mouth around a business and its products and services. In a nutshell, anything published on the SocMed can be seen by millions of people in a very short time.

Despite the numerous advantages of SocMed, its adoption and utilization especially in the contexts of SMEs is not without barriers and problems [5]. Rogova and Prenaj [4] stated that depending on the business type, size, and age and management style of the SMEs, hard efforts need to be made in some areas which include the need to engage human and time resources to manage the SocMed presence, the need to be very active and produce new content regularly so as to stay in the radar of the consumers, the need to control the contents to be published so as to avoid any reverse effects on the SMEs image and reputation. The aforementioned challenges are among the reasons why SMEs are still not taking full advantage of the SM presence.

As opposed to SocMed and CRM, companies have also widely used them for establishing contacts with their customers. The implementation of CRM has shown many positive impacts to the company's wellbeing which include increasing organizational performance, improving revenue and profits [6, 7]. Despite the positive testimonies of CRM implementation, there were also reports on its problems and failures. For instance [8] found that 70% CRM projects resulted in loss and showed no improvement due to lack of knowledge and financial resources.

Mining the literature suggests that studies on SocMed and its integration into CRM by SMEs is still very limited. Furthermore, the available literature was mainly done in countries outside Malaysia. While the findings of these studies are undoubtedly helped the researchers to better understand SocMed utilizations, they are however not easily applied or implemented in the Malaysian settings. To this effect, a study that focuses on the use of SocMed and its integration into CRM among SMEs in Malaysia is considered crucial.

2 Literature Review

The study adopts and adapts a few dimensions to measure business performance, including operational performance, customer satisfaction, market effectiveness, and profitability/financial performance [9]. Meanwhile social media pricing capability, social media product development capability, and social media marketing are dimensions measuring social media [10]. For Customer Relationship Management (CRM), the dimensions are customer relationship orientation, customer-centric management system, and relational information processes [11].

2.1 Business Performance

Sustainability of an organization is highly depending on their business performance. There are few factors that contribute to business performance, i.e., operational performance, customer satisfaction, market effectiveness, and profitability/financial performance [9]. Organizations have more control on operational performance as it is less affected by external factors. Customer satisfaction, market share, reduction in management cost, lead and order time, effective usage of raw material as well as enhancing the production activities' effectiveness [12, 13] are among the components that led to operational performance which will increase the business performance. Second dimension in business performance is customer satisfaction. Customer loyalty, purchase intentions, and organizations care, i.e., appreciate the customers' time and effort are among the most significant indicators of customer satisfaction [14]. In addition, the use of social media is seen as a perfect opportunity for organization to communicate with customers and solve issues within short period of time [15]. Marketing effectiveness is essential in measuring the business performance. According to [16], marketing effectiveness is where organizations appraise their marketing activities and strategies in getting their consumer's or customer's preferences, needs, and satisfaction. There are several ways for organizations to put marketing effectiveness in place such as perform corporate tasks, enhance corporate performance, interact socially, depend on each other to complete tasks, and build profitable strategies for competitive advantage [17–19]. The fourth dimension measuring business performance is profitable/financial performance. The competitive pressure resulting from market interactions between entrants and incumbents plays a significant role in determining the stability differentials in profits. It also affects the balance between growth and exit events, cost reduction efforts, margin gap, and prices set by customers [15, 18]. Engaging in these activities can create new opportunities to boost the organization's profitability and efficiency.

2.2 Social Media (SocMed)

SocMed is claimed to be very convenient for customers, as they can browse and purchase easily with minimal effort. In this study, three dimensions have been used to measure SocMed, such as social media pricing capability, social media product development capability and social media marketing [10]. Back in 2011, 39% of organizations used social media as primary digital tool in order to reach customers effectively and expected to increase to 47% within the next four years [20]. Organizations believe that with the use of SocMed, they can connect and engage their customers, reach larger and broad media expansion with high efficiency and low costs compared to traditional media [21, 22]. Organizations may be able to hit their potential customers based on psychographic and demographic characteristics with the features provided by SocMed too [23]. With that, organizations can use the advantages of SocMed to become one of their sales and marketing platforms, enhance their product/service development capabilities, create engaging content, gather customer's reviews and much more. One of the benefits in using SocMed among organizations is products/services development, for example, PepsiCo learns from their potential customers in creating new varieties of products and has sold more than 36 million cases of the Mountain Dew brand products [24]. To one extend, organizations have more essential to competitive advantage than others [25]. According to [26], SocMed is designed to be an interactive platform for their users and communities to co-create, share, modify, and discuss user-generated content. All these led to the rose number of SocMed users from year to year, i.e., 2022 (up to October) with 59.3% and 2021 (up to October) is about 57.6% out of world population with 1.7% increment [1]. To organizations, the increment number of SocMed users shows that this platform is a right platform to market their products/services apart from gaining more customers and sales. Nevertheless, there are some organizations that deny the capabilities of SocMed, refuse to use it, and gain benefits it offers [24, 26].

2.3 Customer Relationship Management (CRM)

Customer relationship management is an Information System (IS) application for capturing customer's profile. Customer relationship orientation, customer-centric management system, and relational information processes are the dimensions used in this study [11]. Customer relationship orientation can be defined as the capability of an organization to build, develop, and maintain a relationship with the customer through which the organization achieves its business goals, ensures its survival, and builds resources itself by nurturing relevant relationships with its stakeholders [27]. In addition, the benefits of having customer relationship orientation implementation are getting profitable customers' needs, assisting partner strategies that may influence customer resources, and also improving an organization's efficient and effectiveness [28]. In ensuring to achieve the customer relationship orientation benefits, organization can incorporate technology in their marketing activities besides produce new ways of undertaking business as well as enhancing the relationships with customers [29]. Apart from that, organizations that use SocMed are able to respond to their customers' concerns/insight/complains in a prompt manner. There is no doubt that SocMed is a new method of communication between customers and organizations [29]. Whenever CRM is in place, organizations should have a complete set of their customers' information, behavior, liking, and beyond. As a result, developing a customer-centric relationship management system can pose challenges for organizations in gathering and managing such information [30]. An effective and efficient CRM could assist organizations beyond ordinary business processes and able to fulfill their customer requirements and needs [31]. These capabilities are necessary for organizations to create customer-centric management system for them to understand their customers better. Consequently, the failure to link what their customers wants and needs will lead and cost the organizations to spend more time and money. It is well-known that CRM is a part of organizational process that focuses on establishing, enhancing, and maintaining long-term associations with their customers [11]. Nevertheless, the information processes are likely to be influenced by an organization's management system. The availability of CRM enables the organizations to use it for establishing and maintaining relationships, as well as providing appropriate responses to customer needs [11]. Therefore, information plays a vital role in creating and maintaining the customer relationships.

3 Research Methodology

This pilot study was carried out in SMEs in Klang Valley, Malaysia, with specific sectors such as food and beverage and textile. This selection sector was chosen based on the most extensive use of SocMed. The questionnaire was distributed in the state of Selangor and followed by Kuala Lumpur, Cyberjaya, and Putrajaya. The researchers distributed around 40 copies of the questionnaire as the pilot study only required a minimum of 30 respondents [32] for further analysis of reliability.

The items in the questionnaire are adapted from the work by Reimann et al. [9] and Tarsakoo and Charoensukmongkol [10] with a few modifications in term of language and content. The important factors in this pilot study were to ensure that all the items in the questionnaire are reliable and correct to address the research objective later on. Furthermore, researchers will like to ensure that the items were clearly understood, well presented, and defined.

Table 1 Demographic profile

		Frequency	Percentage (%)
Year(s) of existence in business	5 years and below	14	35
	6–10 years	6	15
	11–15 years	3	7.5
	16–20 years	3	7.5
	21–25 years	8	20
	> 25 years	6	15
Sector	Food and beverages	35	87.5
	Textile	5	12.5
Year(s) of social media marketing experience in business	5 years and below	24	60
	6–10 years	13	32.5
	11–15 years	1	2.5
	16–20 years	1	2.5
	> 25 years	1	2.5

4 Findings

About 30 respondents' data were entered into the Statistical Package for Social Science (SPSS) version 26. Descriptive statistics for demographic profile and reliability test by looking the Cronbach alpha value in each group of DV and IV are the main analysis that has been done.

4.1 *Demographic Profile*

Table 1 presents the descriptive analysis of the respondent's demographic profile. The majority of the SMEs have a business existence of 5 years and below. 87.5% were involved in the food & beverages sector, while the remaining 12.5% belonged to the textile sector. In term of year(s) of social media marketing experience in business, the majority of SMEs had 5 years and less of experience.

Table 2 indicates the SocMed account maintained by the SMEs. The respondents are allowed to answer more than 1. It is found that Facebook and Instagram are the most SocMed account that they have with 43.2% and 27.3% respectively.

4.2 *Reliability Analysis*

This study has performed reliability analysis in order to assess the scale's internal or reliability consistency strength. Table 3 indicated that all variables are above the

Table 2 Social media account

		Frequency	Percentage (%)
Social media	Friendster	0	0
	Linkedln	5	5.7
	Myspace	0	0
	Facebook	38	43.2
	Twitter	6	6.8
	Instagram	24	27.3
	Snapchat	0	0
	Others	15	17

Table 3 Reliability analysis

Variable	No. of Items	Cronbach alpha
Business performance		
Operational performance	5	0.744
Customer satisfaction	4	0.865
Market effectiveness	4	0.919
Profitability/financial performance	4	0.949
Social media		
Social media pricing capability	5	0.893
Social media product development capability	5	0.853
Social media marketing	8	0.901
Customer relationship management (CRM)		
Customer relationship orientation	4	0.900
Customer-centric management system	6	0.847
Relational information processes	8	0.930

recommended cutoff value of the pilot study which were 0.6. Therefore, the scale used in this study was highly validated [33].

5 Discussion and Conclusion

The protocol approach in this pilot study has shown feasibility. It is clearly shown that the study did not appear to have any problem or impact on the SMEs companies' employees. On top of that, researchers will follow-up with respective respondent to give their feedback either via phone call or email. Above all, the questionnaire does not need any amendments as this pilot study revealed no flaws or issues with the items being assessed or commented on. For researchers, it presents a significant challenge as most companies continue to operate online due to the Pandemic Covid-19 and the only way for data collection is through email. Researchers need to get mobile phone number in order to reach companies under SMEs for follow-up purposes. The reliability analysis shows that all independent variables and dependent variables meet the acceptable cutoff point to proceed to the main study. It is important not to overlook the pilot study process as it provides numerous advantages. Undoubtedly, the pilot study process carries some risks. However, investing resources and time into it is worthwhile because it helps to avoid or eliminate unforeseen difficulties [34].

Acknowledgements This research was funded by Universiti Teknologi MARA (UiTM) Selangor Branch (UCS): DUCS 2.0: 600-UiTMSEL (PI. 5/4) (047/2020).

References

1. Kemp S (2022) Digital 2022 October global statshot report, retrieved from https://datareportal.com/reports/digital-2021-october-global-statshot on 5 November 2022, 2021
2. Trainor JK (2012) Relating social media technologies to performance: a capabilities-based perspective. J Personal Selling Manage 32(3):317–331
3. Guha S, Harrigan P, Soutar G (2017) Linking social media to customer relationship management (CRM): a qualitative study on SMEs. J Small Bus Entrep 30(3):193–214
4. Rogova B, Prenaj B (2016) Social media as marketing tool for SMEs: opportunities and challenges. Academic J Bus Adm Law Soc Sci 2(3):85–97
5. Stelzner AM (2014) 2014 Social media marketing industry: how marketers are using social media to grow their businesses. Soc Med Exam 1–50
6. Ata UZ, Toker A (2012) The effect of customer relationship management adoption in business-to-business markets. J Bus Ind Market 27(6):497–507
7. Mohamad SH, Othman NA, Jabar J, Majid IA (2014) Customer relationship management practices: the impact on organizational performance in SMEs of food manufacturing industry. Eur J Bus Manage 6(13):35–48
8. Richard JE, Thirkell PC et al (2007) An examination of customer relationship management (CRM) technology adaption and its impact on business-to-business customer relationships. Total Qual Manag 18(8):927–945
9. Reimann M, Schilke O, Thomas JS (2009) Customer relationship management and firm performance: the mediating role of business strategy. J Acad Market Sci
10. Tarsakoo P, Charoensukmongkol P (2020) Dimensions of social media marketing capabilities and their contribution to business performance of firms in Thailand. J Asia Bus Stud 14(4):441–461
11. Jayachandran S, Sharma S, Kaufman P, Raman P (2004) The role of relational information processes and technology use in customer relationship management. J Mark 69(4):177–192
12. Azim MD, Ahmed H, Khan ATMS (2015) Operational performance and profitability: an empirical study on the Bangladeshi ceramic companies. Int J Entrepreneurship Dev Stud (IJEDS) 3(1):63–73
13. Truong HQ, Sameiro M, Fernandes AC, Sampaio P, Duong BAT, Duong HH, Vilhenac E (2017) Supply chain management practices and firms' operational performance. Int J Qual Reliab Manage 34(2):176–193
14. Peek S (2022) Make 'em smile: what drives successful customer satisfaction? Business.com. Retrieved from https://www.business.com/articles/what-drives-successful-customer-satisfaction/

15. Cotriss D (2022) Social media for business: marketing, customer service and more. Business News Daily. Retrieved from https://www.businessnewsdaily.com/7832-social-media-for-business.html
16. Mugwati M, Bakunda G (2019) Board gender composition and marketing effectiveness in the female consumer market in Zimbabwe. Gender Manage: An Int J 34(2):94–120
17. Vandewaerde M, Voordeckers W, Lambrechts F (2011) Board team leadership revisited: a conceptual model of shared leadership. J Bus Ethics 104(104)
18. Bottazzi G, Secchi A, Tamagni F (2008) Productivity, profitability and financial performance. Ind Corp Chang 17(4):711–751
19. Singh V, Terjesen S, Vinnicombe S (2008) Newly appointed directors in the board room: how do women and men differ? Eur Manage J 26(1):48–58
20. Davis C, Freundt T (2011) What marketers say about working online. McKinsey Quart
21. Kaplan AM, Haenlein M (2010) Users of the world, unite! the challenges and opportunities of social media. Bus Horiz 53(1):59–68
22. Thackeray R, Neiger BL, Hanson CL, McKenzie JF (2008) Enhancing promotional strategies within social marketing programs: use of web 2.0 social media. Health Promot Pract 9(4):338–343
23. Constantinides E, Fountain SJ (2008) Web 2.0: conceptual foundations and marketing issues. J Direct Data Digit Mark Pract 9(3), 231–244
24. Divol R, Edelman D, Sarrazin H (2012) De-mystifying Social Media. McKinsey Quart 2(12):66–77
25. Edelman D (2010) Branding in the digital age: you're spending your money in all the wrong places. Harv Bus Rev 88(12):62–69
26. Kietzmann JH, Hermkens K, McCarthy IP, Silvestre BS (2011) Social media? get serious! understanding the functional building blocks of social media. Bus Horiz 54(3):241–251
27. Balakrishnan MS (2006) Customer relationship orientation—evolutionary link between market orientation and customer relationship management. In: 6th global conference on business and economics
28. Gruen TW (1997) Relationship marketing—the route to marketing efficiency and effectiveness. Business Horizons, pp 32–38
29. Venciūtė D (2018) Social media marketing—from tool to capability. Manage Organ Syst Res 79(1):131–145
30. Teoh SY, Pan SL (2009) Customer-centric relationship management system development: a generative knowledge integration perspective. J Syst Inf Technol 11(1):4–23
31. Soh C, Sia SK, Yap TJ (2000) Cultural fits and misfits: is ERP a universal solution. Commun ACM 43:47–51
32. Saunders M, Lewis P, Thornhill A (2012) Research methods for business students, 6th edn. Pearson, London, p 266
33. Nunnally JC (1978) Psychometric theory. McGraw-Hill, New York
34. Hassan ZA, Schattner P, Mazza D (2006) Doing pilot study: why is it essential? Malaysian Family Phys 1(2&3):70–73

Machine Learning-Based Intrusion Detection for IOT Devices

Kirti Ameta and S. S. Sarangdevot

Abstract The Internet of Things (IoT) and its many potential uses are now very much popular subject of study. One of IoT's defining features is how readily it can be implemented in practical settings, yet the same trait also makes it vulnerable to cyberattacks. An intrusion is an attack on a system or network by an unauthorized user who poses as a legitimate user or makes advantage of a security hole or flaw. Intrusion Detection Systems (IDS) are designed to identify attacks at several layers. To that end, this study employs machine learning approaches based on decision trees to better identify and categorize intrusion attempts inside an Intrusion Detection System.

Keywords Intrusion detection system · IOT devices · Internet of things · Machine learning

1 Introduction

In order to identify malicious behavior and common threats, the network traffic will be monitored by an Intrusion Detection System (IDS). After detecting suspicious behavior, it may also send notifications to the administrator. Several ML methods may be utilized to effectively manage and categorize threats. Methods for detecting an incursion are discussed in this section. A hardware or software intrusion detection system (IDS) monitors, identifies, and alerts the computer or network in the event of an assault or incursion. To identify and fix any system or network vulnerabilities, administrators and users may consult this alert report. Anomaly detection, signature detection, and hybrid detection are three typical types of intrusion detection.

Different intrusion detection methods exist, such as host-based detection and network-based detection. Host-based intrusion detection is implemented on a single

K. Ameta (✉) · S. S. Sarangdevot
Department of Computer Science and Information Technology, JRN Rajasthan Vidyapeeth (Deemed to Be) University, Udaipur, Rajasthan, India
e-mail: kirti09pune@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_81

host and used to keep tabs on all incoming and outgoing data and compare it to a model of the host's traffic flows. This often involves a software agent that monitors the host's activity and looks for signs of infiltration by examining things like system calls, application logs, directory changes, and other user actions. In order to identify malicious behavior, network-based intrusion detection analyzes network traffic and keeps tabs on numerous hosts inside the network. Captured network traffic is analyzed at the network, transport, application, and hardware layers in an effort to unearth malicious behavior.

Machine Learning (ML) is a branch of Artificial Intelligence (AI). Machine learning enables systems to acquire and hone skills automatically via exposure to new data and use, all without requiring human intervention. The ML method is more effective for IDS in identifying assaults for large amounts of data in a shorter period of time. Generally speaking, ML algorithms may be broken down into one of three types: Supervised, Unsupervised, and Semi-Supervised.

The supervised algorithm investigates information that has been completely class labeled. You may use either classification or regression to do this. Training and testing are two phases of the categorization process. The response variable serves as a data aid throughout the training process. Classification techniques include the usual suspects like the Support Vector Machine (SVM), Discriminant Analysis, Nearest Neighbor, Naive Bayes, ANN, and Logistic Regression (LR). Linear Regression, SVR, Ensemble Methods, Decision Tree, and RF are all examples of algorithms that fall within the broader area of regression.

2 Background

With IDS in place, organization will have a secure environment in which they can carry out their activities, and it will be able to prevent harmful network intrusions. IDS systems today often use Machine Learning (ML) methods as a means of improving their ability to identify and classify potential security risks. In this study, we evaluate the use of a variety of ML approaches in IDS and compare their respective levels of effectiveness.

Zhong et al. [8] revealed that it has been widely discussed how crucial a part IoT plays in people's everyday lives, with orders and data transferring quickly between computers and things to facilitate service delivery. But cyberattacks have become a major concern, particularly for servers used in the Internet of Things. Network backbones need to be fortified against a wide range of threats. The Intrusion Detection System (IDS) serves as the unseen protector of IoT servers. Intrusion detection systems (IDS) have made extensive use of machine learning techniques. Even still, the IDS system might need some refinement in terms of precision and efficiency.

Alsoufi et al. [1] claim that the use of deep learning algorithms as a method for preserving the environment of the internet of things has been successful. The widespread use of deep learning as a defense mechanism against intrusion detection attempts is evidence of the effectiveness of this method. IDS that are based on anomalies, rather than signatures, are more able to spot zero-day attacks than signature-based systems.

Liang et al. [4] reports have shown that conventional IDS do not fare well in the IoT's network environment; hence, research on intrusion detection systems well-suited to the IoT's network environment is warranted. Researchers have discovered that integrating machine learning technologies into an IDS is an efficient solution to address the limitations of conventional IDSs in the context of the IoT. Their study includes developing and testing a new model of analysis for use in intrusion detection systems. In order to identify intrusions, the latest system employs a multi-agent-based hybrid placement method. The new system is divided into four sections: data gathering, data management, analysis, and reaction.

Smys et al. [7] An IoT network intrusion detection system based on a hybrid convolutional ANN model was suggested in this study. The proposed paradigm may be used in a variety of Internet of Things contexts. As of 2016, Hodo et al. have presented an offline IDS for IoT. The device uses an ANN to sift through IoT network data and spot DDoS assaults (ANN). Their deployment plan centers on keeping an eye on the data flowing via the Internet of Things in order to identify DoS and DDoS assaults. In an IoT setting, all packet traffic is monitored by a single, centralized system [3].

In 2018, researchers Diro and Chilamkurti suggested a deep learning-based distributed IDS for IoT. The detection system relies on an anomaly detection technique. In comparison to conventional IDSs, their investigations show that IoT/Fog network attacks may be detected with high accuracy [6]. While intrusion detection technologies are commonplace in traditional networks, not much thought has been given to how machine learning may improve IoT security. To the contrary, as Raza et al. point out, typical intrusion detection systems are not sufficient to safeguard IPv6-connected IoT or more sophisticated IoT network settings. Few researches have looked at the potential of machine learning technology to enhance Internet of Things intrusion detection systems [2, 5].

3 Methodology

Data was gathered, and it did include network activity. The KDD dataset not only helped researchers better understand various intrusion patterns, but is also frequently used as a benchmark for assessing the efficacy of various intrusion detection tools. In order to derive fair experiments from this dataset, statistical measures give a thorough knowledge of the dataset itself. The attacks in this dataset may be roughly classified into four classes:

1. Denial of Service (DoS): Synonymous with causing a host or server to crash, these assaults are disruptive to routine operations. Such an assault occurs when, the attacker either denies access to genuine users or causes memory exhaustion, preventing the system from processing requests from valid users.

2. R2L: Remote to Local (R2L) refers to the situation in which an attacker may bypass regular authentication and take control of a system from another location, for as by guessing a password. In such an assault, an attacker may get unauthorized access to a system by sending a package across a network without actually logging in as a user to the computer.
3. User to Root (U2R): As the name implies, it includes an adversary posing as the network's administrator by stealing credentials from a trusted user. Attempts to get access to the local superuser (root) account without permission, such as different "buffer overflow" assaults. In this scenario, the attacker compromises a normal user's account by obtaining access to that user's credentials and then using those credentials to gain control of the system.
4. Probe: In this kind of assault, data is gathered in preparation for a future incursion. This may involve surveillance activities in addition to another kind of probing attack, such as port scanning. An attacker gathers data about a computer system network with the intention of discovering a technique to bypass security measures.

In most instances, the data that is used to generate the final model is a compilation from a number of various sources. This is because the data that is used to develop the model is the most accurate. It is necessary to initially divide the data into a training set and a testing set before attempting to learn anything from them. The parameters of the model are tailored to a selection of cases included within the training dataset, and the model's functionality is then objectively assessed based on how well it performs on an independent test dataset. This research follows a well-known approach that recommends separating data into training and testing sets with a ratio of seventy percent training and thirty percent testing.

4 Results

Code:

```
# Apply machine learning classification algorithms Decision Tree
To use, just import sklearn.tree from your working directory.
DecisionTreeClassifier
useplt for import of matplotlib.pyplot
The Tree Importer from Sklearn
fromsklearn.tree import export_text
clfd = DecisionTreeClassifier(criterion = "entropy", max_depth = 4)
start_time = time.time()
clfd.fit(X_train, y_train.values.ravel())
end_time = time.time()
print("Training time:", end_time-start_time)
start_time = time.time()
y_test_pred = clfd.predict(X_train)
```

```
end_time = time.time()
print("Testing time:", end_time-start_time)
print("Trainscore is:", clfd.score
(X_train,y_train))
print("Testscore is:", clfd.score
(X_test,y_test))
plt.figure(figsize = (20,20))
#create the tree plot
a = tree.plot_tree(clfd, rounded = True, filled = True, fontsize = 16)
#show the plot
plt.show()
#text based diagram
#export the decision rules
tree_rules = export_text(clfd)
#print the result
print(tree_rules)
```

See Fig. 1.

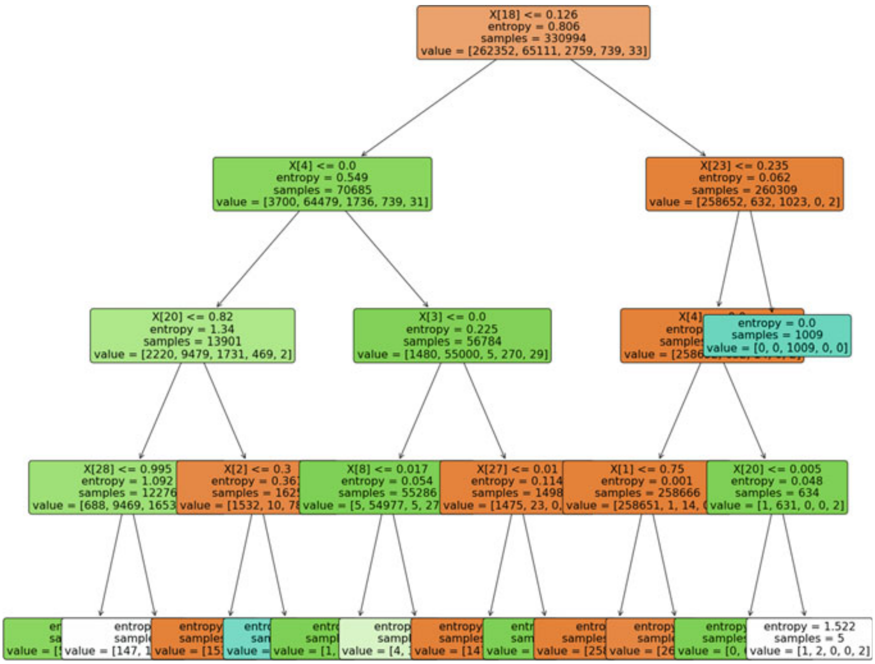


Fig. 1 Decision tree diagram for IDS for KDD data set.

5 Conclusion and Future Scope

A decision tree is a method of classification that uses a series of choices, each of which contributes to the next. A tree structure may be used to illustrate such a series of choices. Classifying a sample involves working its way outward from the root node to the appropriate end leaf node, with each leaf node standing in for a different category. Decision trees are used to forecast the value of a target class for an unseen test instance based on the values of numerous known examples (DT). A decision tree is a classification method that uses a series of judgments to determine how an unseen test case should be labeled. Since decision trees are so straightforward and easy to build, they are often used as a single classifier.

Decision tree algorithms work by connecting inputs with predetermined outcomes. Accordingly, we use a certain set of inputs to get to the output. Statistics, data mining, machine learning, and other disciplines often make use of this kind of modeling. Classification trees are a kind of decision tree in which the "leaves" represent the target variables.

Decision tree categorization seeks to provide information in a structure that includes both the root and the leaf nodes. In order to detect malicious actions, decision trees may analyze data and identify key characteristics in the system. By verifying the order of intrusion identifiers, this boosts the effectiveness of several security systems. It can recognize instances and patterns that encourage checking, advance attack signatures, and identify a variety of checking actions. Compared to alternative approaches, using a decision tree provides a wealth of rules that are both simple and intuitive, and it integrates well with real-time technology.

References

1. Alsoufi MA, Razak S, Siraj MM, Nafea I, Ghaleb FA, Saeed F, Nasser M (2021) Anomaly-based intrusion detection systems in iot using deep learning: A systematic literature review. Appl Sci 11(18):8383
2. Al-Yaseen WL, Othman ZA, Nazri MZA (2017) Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system. Expert Syst Appl 67:296–303
3. Asharf J, Moustafa N, Khurshid H, Debie E, Haider W, Wahab A (2020) A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions. Electronics 9(7):1177
4. Liang C, Shanmugam B, Azam S, Jonkman M, De Boer F, Narayansamy G (2019) Intrusion detection system for Internet of Things based on a machine learning approach. In: 2019 International conference on vision towards emerging trends in communication and networking (ViTECoN). IEEE, pp 1–6
5. Lin WC, Ke SW, Tsai CF (2015) CANN: an intrusion detection system based on combining cluster centers and nearest neighbors. Knowl-Based Syst 78:13–21
6. Mandal K, Rajkumar M, Ezhumalai P, Jayakumar D, Yuvarani R (2020) Improved security using machine learning for IoT intrusion detection system. Mater Today: Proc
7. Smys S, Basar A, Wang H (2020) Hybrid intrusion detection system for internet of things (IoT). J ISMAC 2(04):190–199

8. Zhong M, Zhou Y, Chen G (2021) Sequential model based intrusion detection system for IoT servers using deep learning methods. Sensors 21(4):1113

Seek N Book: A Web Application for Seeking Gigs and Booking Performers

Eric Blancaflor, Jeanne Bernaldo, Elijah Lowell Calip, and Pauline Andrea Vivero

Abstract The Seek N Book web application, designed for seeking gigs and booking performers, has the potential to make a significant contribution to the music society. It can greatly assist musically inclined people in securing jobs within the music industry, showcasing their talents, and fostering connections among musicians, organizers, and fans. This is particularly relevant due to the growing influence of various music genres facilitated by social media and other platforms. Previous studies show that musicians were able to use the web application although interactions are very limited. The proposed web application in this study enables the users, which is the organizer and the musician, to create a job and communicate with each other. The design shows its effectiveness in creating a platform specifically in connecting organizers and musicians.

Keywords Web applications · Gigs · Booking · Musician · Performers

1 Introduction

1.1 Project Context

Music has been a way of expressing oneself and can be a source of their livelihood. Either play in a band, become a member of the orchestra, create songs, and make music for movies, television shows, or commercials. Every event around the world has musicians play music to their audience. Implementing an event is not separated from the services of musicians, let alone a music event, music festivals, and other events. Musicians predominantly dominate the provision of services for events, but the scarcity of essential information makes it arduous to uncover their existence and

E. Blancaflor (✉) · J. Bernaldo · E. L. Calip · P. A. Vivero
School of Information Technology, Mapúa University, Manila, Philippines
e-mail: ebblancaflor@mapua.edu.ph

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_82

connect with them [1]. It will undoubtedly take a long time for musician seekers to find musicians to make a booking.

In the Philippines, it is widely acknowledged that there are outstanding performers, and some even receive recognition on both local and international stages. However, interesting issues persist for this generation of music players and downloaders. Despite the recent advancements in technology, a significant number of musicians continue to be left behind, commonly known as Unsigned Musician. These talented individuals possess exceptional compositions and music that have the potential to contribute greatly to the music industry, yet they remain unheard. The aforementioned factors contribute significantly to the explanation as to why nine out of ten newly signed musicians fail to record, let alone release, a second record [2].

Today, as much as they want to make their music known at its best quality, it is being surpassed by what the popular music industries promote. Filipino masses are innately attracted to anyone or anything famous, especially when seen in their favorite movies and soap operas [3] and in addition, the rise of new digital platforms has not only enabled new forms of work activity but has also fundamentally transformed the way freelancers find new opportunities [4]. For these unsigned musicians to gain popularity or have the opportunity for their career, they must look for gigs or do any freelancing job to somehow keep up with the cost of living and aid their needs to fulfill their passion which is conquering the music industry.

1.2 Purpose and Description of the Study

The proposed web application enables the users, which is the organizer and the musician, to create a job and communicate with each other with the listeners as a default user:

The project can be summarized from the perspective of each user:

- User1(Organizer): The main feature for User1 is that they can book musicians for their events. The organizers can also create a job wherein musicians or listeners can apply.
- User2(Musicians): Musicians are the only users who can receive a booking request. They can accept or decline a booking message/request from an organizer. The musician can also create a job wherein organizers or listeners can apply.
- User3(Listeners): These users can apply as band members or even qualify as musicians to any organizer's event that can also view and react to posts. These users can use some of the web's functionalities but creating an event or job is not applicable.

This chapter aims to help determine the current situation of local unsigned musicians through interviews and develop a web application to provide a more straightforward booking and job finding of their services.

1.3 Scope and Delimitations

The main users of the web application will only include local unsigned musicians, event organizers, and music listeners. Only event organizers are allowed to book musicians and create events and not vice versa. Both musicians and organizers can create a job. All users can create threads, create posts, react to posts, send messages, and send an audio file. Registered users may only access the web application. It will help assist any musicians in forming bands and promoting their activities.

This project is specifically intended for local unsigned musicians having difficulties finding gigs around the city and event organizers looking for suitable musicians for their events. The project aims to make booking much more straightforward than the old way, which takes time to find a gig or book musicians. The project will not focus on creating contracts, processing payments, and page group creation, and some features might not work on mobile phones.

1.4 Significance of the Study

This project benefits local unsigned musicians (solo artists, bands, and other musicians) and organizers. The project would also be a valuable tool to assist future students who would need additional resources to study booking systems for local musicians.

2 Review of Related Literature

2.1 Web Application and Its Components

A Web application is an application that is invoked with a web browser over the Internet. Ever since the 90s when the Internet became available to the public the World Wide Web put a usable face on the Internet, the Internet has become a platform of choice for many ever-more sophisticated and innovative Web applications. In just one decade, the Web has evolved from being a repository of pages used primarily for accessing static, mostly scientific, information to a powerful platform for application development and deployment [5]. Additionally, web applications do not rely on the operating system of your desktop as it is accessed through internet browsers such as Google Chrome or Mozilla Firefox which run on web servers able to display information regardless of the operating system of your desktop computer. Web applications also store the information provided by their users remotely as opposed to desktop applications which store information on the user's computer [6].

Web applications are composed of a set of hardware and a collection of scripts and programming languages. For the client-side, the scripting language used are combinations of HTML which is used to structure a webpage and its content, CSS to design and style documents on the web, and JavaScript to control the behavior of different elements. Using these together allows for dynamic content on the client-side and is responsible for everything a user sees on their browser. For the "back-end" or the server-side, there are scripting and programming languages available which allow a web application to fulfill its purpose of serving dynamic content to users. PHP can be used in server-side scripting, command-line scripting, and writing desktop applications. Through server-side scripting, the web server establishes a connection with the web browser in order to execute the program that the user is trying to run [7]. There are types of databases to store information needed for the web applications, and examples are Microsoft SQL Server, MySQL, and MongoDB.

2.2 PHP MVC Framework

PHP MVC is an application design pattern that separates the data access and business logic (model) from the data presentation (view). PHP is easy to set up, compiles fast, cross-platform, and open-source. Through PHP, developers can maximize what they can include on their websites. Users can have their data stored in the database. These features are beneficial for the development of this project.

Since reliability, scalability, security, and maintainability are all in demand for web-based applications, coding languages such as PHP are becoming popular. However, features such as data access, business logic, and data representation are all put together in one, causing some problems in big projects. To solve this, web applications would make use of the MVC design pattern [8].

The Model View Controller (MVC) design pattern separates different parts of the code into three components, making coding more manageable. As the name suggests, MVC consists of a Model, View, and Controller. The Model component serves as the permanent storage of data. It serves as the bridge between the View and the Controller. Its purpose is to allow the writing and reading of data in the application. However, this component does not consider what happens to the data as it only processes them when needed. The View component allows direct user interaction with the application. It handles the HTML part of the web application, displaying the output of the data processed by the Model from the Controller. The Controller component handles the data submitted by the user then passes it to the Model for storage and updates the Model accordingly when modified. It does not process the data as it is the job of the Model. The Controller only acts when the user interacts with the View [8]. Selecting the best PHP framework is a challenge and some of the commonly used frameworks are CodeIgniter, Kohana, CakePHP, and more.

2.3 Freelancing Systems and Booking App

Based on researcher Thabassum (2013), most workers are always online and in touch with their respective clients. Technological innovations such as electronic deliverability of jobs and fast Internet connections have increased the supply of remote jobs that can be performed by freelancers [9].

Internet booking applications have been viral [10]. These applications have high-concurrence features and perform real-time, high-reliability, and security. According to Xing et al. searching and booking are the fundamental operations that heavily influence system performance. The system must be able to perform its best in these fields.

In this chapter, researchers Fiorentini et al. (2015) discuss the importance of providing a system, method, or process that may allow performers, promoters, and venue owners to negotiate booking contracts through the convenience of [11] electronic Commerce.

They have developed a website that offers a confidential chat forum on a booking website, specifically designed for performers and performance seekers. This platform allows both the performer and performance seeker to view each other's profiles. The performance seeker and performer agree to an electronic contract for a live performance by a performer. Performance seeker or performer records comments about the other party in their profile [11].

The researchers concluded it as an object of their present invention to overcome or ameliorate at least one of the prior art's disadvantages and provide a helpful alternative [11].

Batubara and Bachtiar proposed a booking application of music services for mobile devices to help musicians make the search and ordering process easier by leveraging technological advancements. The built-in app will provide a service that connects musicians with musicians. Musicians can search and book musicians equipped with musician recommendation features based on his portfolio videos utilizing the YouTube API. Search can be done by specifying the desired criteria. For example, users can search for musicians based on their desired genre of music. After that, the system will look for a musician that matches the desired criteria [1].

Moreover, Ben Shneiderman, in his study, distilled the vast corpus of user interface design into handful principles. These principles, derived from experience and refined over three decades, require validation and tuning for specific design domains [12].

3 Methodology

The researchers employed the Software Development model, as seen in Fig. 1, a methodology that makes the researchers to create a system for a technology-based project, ensuring well-defined objectives are established. The Figure below shows how the research and development model is broken down.

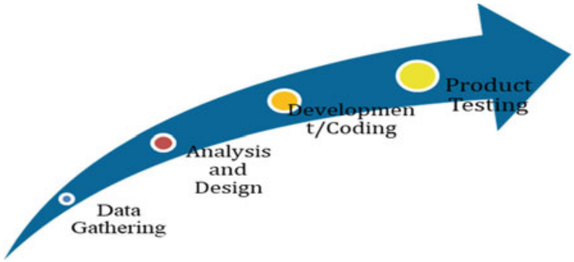


Fig. 1 Software development model

3.1 *Data Gathering Phase*

The first step in this model is data gathering. Here, researchers gather pieces of information and related systems, which also include conducting interviews with specific users (organizers and musicians) of the proposed system. The proposed study has the following features below:

- Recruitment—(a) any user can post a job or look for a job, (b) interested applicants can message the post's author by clicking the job posted, (c) agreements on both sides will stay confidential.
- Response—(a) once an organizer creates an event and book selected musicians, the system automatically sends a message to the invited musician, (b) the musician will choose to accept or decline the organizer's invitation via message, (c) once any user applies for a job, the system will automatically send a message to the creator of the job offer, (d) all users can create a thread for discussion onto the community page and comment to any users' thread, (e) all users can like on other users' posts from their profile, (f) all users can send an audio file, (g) all users can follow each other.
- Booking—(a) must create an event before booking a musician, (b) only organizers can make the booking, (c) organizers shall be the ones providing the contract.

3.2 *Analysis and Design*

This phase includes the breaking down of deliverables into more detailed requirements. The researchers created the process flow for the system and a use-case diagram on how the system will be used, see Fig. 3.

As seen on Fig. 2, specific users (organizers and musicians) can post a job and can be viewed and applied by any other user. Once a user clicked the job offer, it will redirect to a conversation with the author of the offer to discuss further details and may decline if they are not interested or accept the applicant.

As seen in Fig. 3, organizers must create an event before booking a musician. After creating an event, organizers can look for their preferred musicians and click the "Book Now" from the musicians' profiles. Once a musician fully accepts the

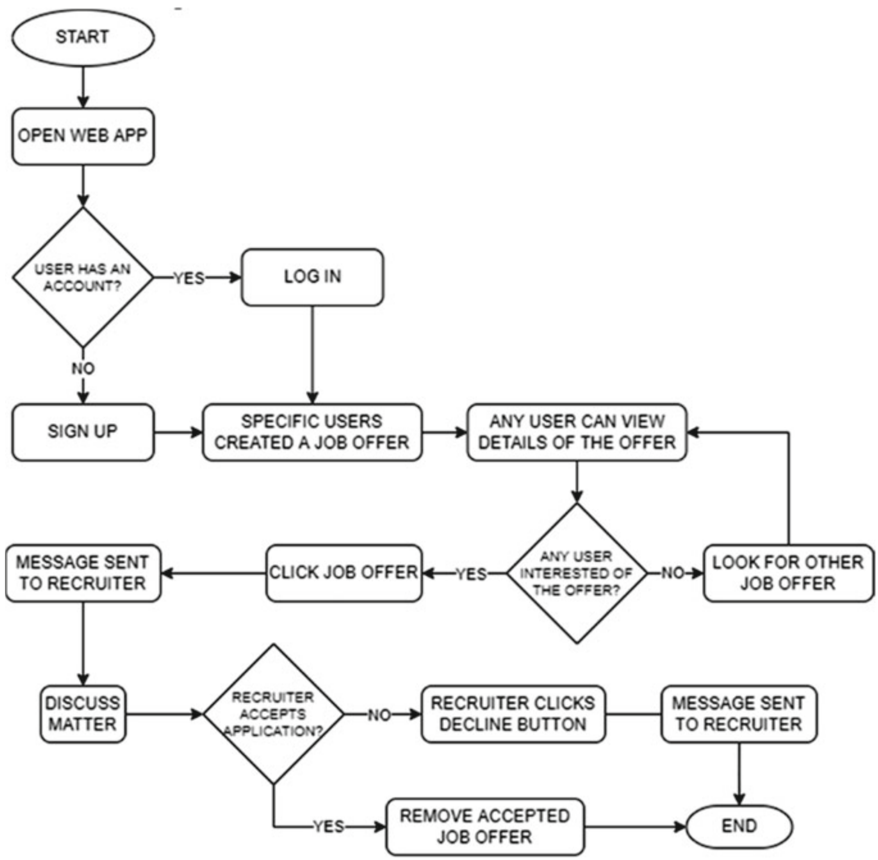


Fig. 2 Recruitment process flow

offer, they can now click the "Accept" button from the organizer's invitation, and their profiles will be added to the organizer's event.

3.3 Development/Coding

This is where the actual coding and implementation of the system takes place. The design will now be turned into code using the programming language decided in the analysis and design phase. The system architecture shown in Fig. 4 presents the hardware and software module design of the proposed system. It is divided into four major components, local application logic, server application logic, database, and external resources.

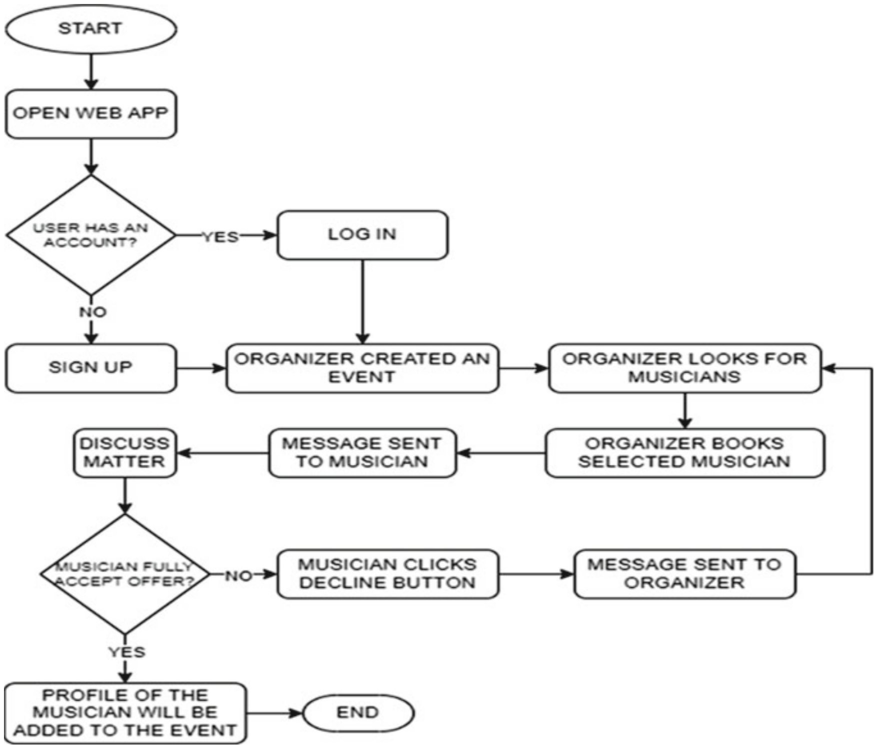


Fig. 3 Organizer booking process flow

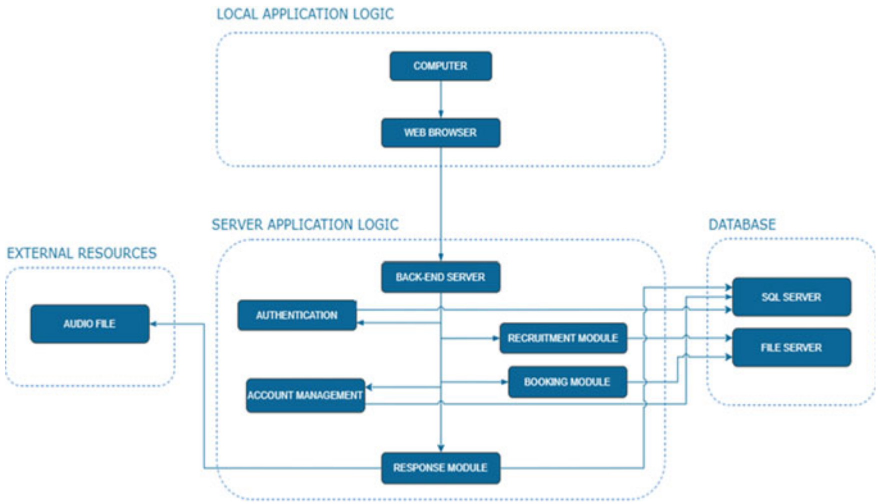


Fig. 4 System architecture

3.4 Product Testing

The product testing phase is where the system and application will be tested out. UAT results conducted in this study are presented in the next section.

4 Results and Discussion

A user acceptance test questionnaire using PSSUQ (Post-Study System Usability Questionnaire) was used to conduct a summative assessment since the review took place after the system had been completed and has undergone testing. The seven-point Likert scales shown in Fig. 6 were used to interpret the data that were gathered from the software evaluation. The questions are divided into three sub-scales, namely System Usefulness (SYSUSE), Information Quality (INFOQUAL), and Interface Quality (INTERQUAL).

The purpose of the initial questions is to identify whether the researcher's developed system met their objectives. This section includes questions that pertain to the user's system experiences.

As presented in Fig. 5, the result of the first sub-scale, System Usefulness (SYSUSE) evaluation, indicates that the users are satisfied using the system. This actively illustrates that the users were able to learn the system at once which enables them to accomplish the tasks and scenarios readily and comfortably.

The next half of the questions is to identify whether the information shown on the website is helping users to find what they are looking for: gigs, events, trending, and so on.

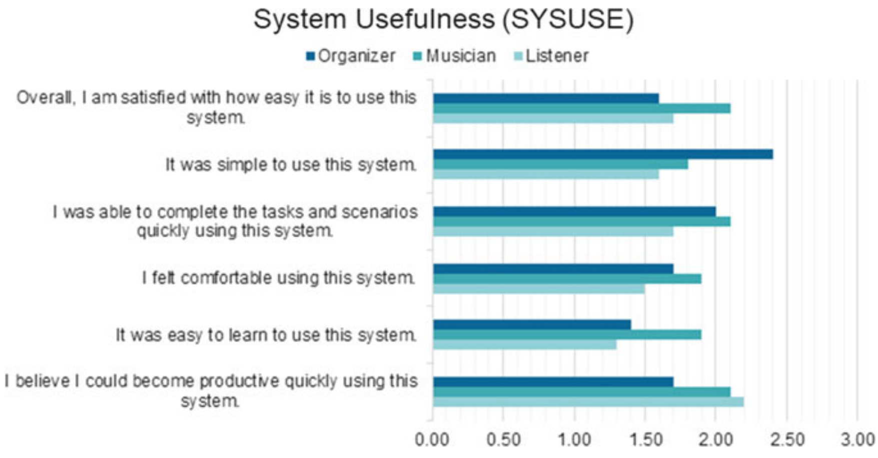


Fig. 5 System usefulness chart

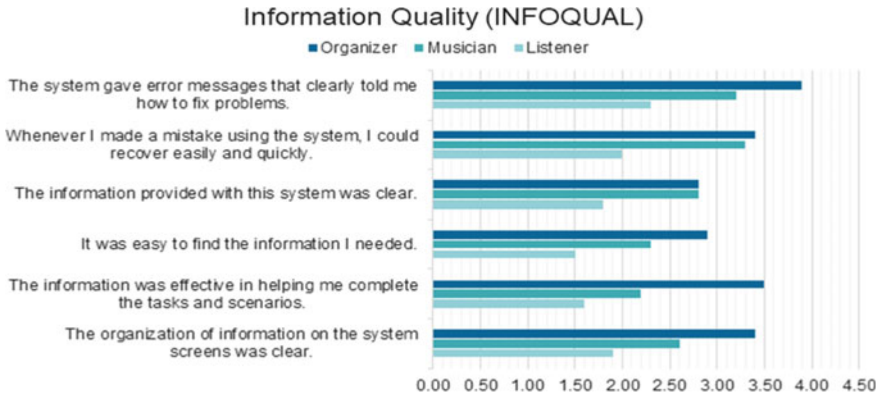


Fig. 6 Information quality chart

The target users may need instructions to perform tasks, the reason for this occurrence is that they were confused about how the system works at their first experience. With instructions, users can comfortably navigate and correct any mistakes they might experience such as editing event information, unfollowing other users, and so on as shown in Fig. 6. Therefore, the users were able to find the information they needed and have given the second sub-scale: System Quality Information (SYSQUAL) a satisfactory rating.

5 Conclusion and Recommendation

Organized information or details play a considerable portion in the decision-making process because better comprehension can be obtained from having a clear view of elements rather than from a visual that is in plain text and numbers. This chapter aims to provide a more straightforward booking and job finding of each users' services. With the researchers' testing, it was observed that the following features seemed to help the target users: organizers and musicians on their process on booking musicians and seeking gigs. For the organizers, the feature of booking musicians proved to be useful as they will just go to the musicians' profiles, send an invite for booking, and wait for their response and once those musicians accepted their requests, they are automatically added to their event. For the musicians, the features of creating a job to find bandmates or organizers proved to be helpful as they will just go create a job or apply for a job and wait for a response. Posting in the community page seemed to be helpful as well, as they can share what they want and have interactions with other users.

Based on the result of the user acceptance testing, it proves that the functionalities were approved by the users with a satisfactory rating to a highly satisfactory rating in accordance with the PSSUQ criteria. With these results, researchers can conclude that the website is proven useful and of great benefit by the end users.

Furthermore, the web application shows its effectiveness by creating a platform specifically designed to connect organizers and musicians. Although there are a few limitations when it comes to the user interface of the website, users were still able to navigate the website. This study could also serve as a guide for future researchers who are planning to conduct this for the local music industry. It is also recommended that future researchers to also give opportunities to other local performers like comedians, stage performers, and so on.

References

1. Batubara R, Bachtiar AM (2019) Booking application music services with YouTube API and GPS sensors based on android
2. Spellman P (2000) Short-term corporate profits versus long-term music careers. In: The self-promoting musician: strategies for independent music success, Berklee Press, pp 8–9
3. Canto JC (2019) Louder for the people in the back,' indie versus popular Filipino music. SunStar, Cebu, 2019. Available: https://www.sunstar.com.ph/article/1798516/cebu/lifestyle/lsquolouder-for-the-people-in-the-backrsquo-indie-vs-popular-filipino-music
4. Sutherland W, Jarrahi MH, Dunn M, Nelson SB (2020) Work precarity and gig literacies in online freelancing. Work Employ Soc 34(3):457–475
5. Jazayeri M (2007) Some trends in web application development. In: Future of software engineering (FOSE '07)
6. TechTerms, "Web Application Definition," Sharpened Productions, February 17 2014. [Online]. Available: https://techterms.com/definition/web_application
7. What can PHP do? The PHP Group, [Online]. Available: https://www.php.net/manual/en/intro-whatcando.php. [Accessed 2021]
8. Olanrewaju R, Ali N, Islam T (2015) An empirical study of the evolution of PHP MVC framework. In: Department of electrical and computer, Kulliyyah of Engineering, Kuala Lumpur
9. Thabassum NF (2013) A study on the freelancing remote job websites. Int J Bus Res Manage 4(1)
10. Zhang Y, Zhao J, Xing C (2009) An extensible framework for internet booking application based on rule engine. In: Sixth web information systems and applications conference, Xuzhou, Jiangsu, China
11. Fiorentini DR, Andrews T, Pollers R (2015) System, process and method of booking musicians and artists. United States Patent 433,345, 22 October 2015
12. Shneiderman, "Golden Rules," University of Maryland, [Online]. Available: https://www.cs.umd.edu/users/ben/goldenrules.html.

Proposal Architecture of the Smart Campus

Salmah Mousbah Zeed Mohammed

Abstract As a high-degree shape of a clever schooling system, the smart campus has received developing research hobby worldwide huge. Due to the multidisciplinary nature of the smart campus, the present studies in the main focus one-sidedly on modern-day technology or revolutionary academic standards, however, leaves a deep perception of the fusion of them and omit the implication of the clever campus in different clever towns. This examination underlines the interdisciplinary imagination and prescient of the clever campus. Based on a complete overview of permitting technology and present clever campus proposals, a human-centric, gaining knowledge of orientated clever campus is envisioned, described, and designed, with the number one purpose of serving stakeholder hobbies and academic transport on the tempo of technological improvement to growth. As properly as discussing the interdisciplinary elements using or limiting the clever campus revolution. The anticipated contribution of this exam is to provide a comparative reference of a smart campus for international schooling providers, governments, and generation organizations supplying such offerings.

Keywords Cloud computing · Internet of things · Augmented reality · Artificial intelligence

1 Introduction

With the improvement of the era, people's dwelling and operating behavior in addition to gaining knowledge of techniques has changed significantly. The slow extrusion inside the gaining knowledge of surroundings and the growing call for personalized and adaptive gaining knowledge of have pushed reforms and tendencies in schooling. As a high-stage shape of a smart schooling system, the clever campus has emerged as a truth and is receiving extra and extra interest across the world. The clever

S. M. Z. Mohammed (✉)
The School of Computer Sciences, Sirte University, Sirte, Libya
e-mail: salmah19@su.edu.ly; salmah.mousbah@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_83

campus creates a smart gaining knowledge of surroundings for residents by reworking them into a smart workforce, thereby turning into a vital part of the clever town framework [1]. The improvement and recognition of smart campuses additionally aid the knowledge-primarily-based totally economy. The global clever schooling marketplace is projected to develop at a CAGR of 15.96% from 2018 to 2022 [2]. In any such hastily changing field, there may be a pressing want to behavior lively studies and apprehend the clever campus and its traits clearly.

Various literature evaluations had been performed on this area concerning the multidisciplinary nature of clever campus studies. On the only hand, the latest emergence and development of records and conversation technology (ICT), synthetic intelligence, clever devices, and variable truth technology are growing unheard-of and forward-looking possibilities for instructional establishments to acquire higher instructional requirements and achievements aid review articles with inside the era segment emphasize cutting-edge technology and search for their capacity programs on the clever campus. To call only a few, Internet of things (IoT) programs and cloud computing technology with inside the clever campus are mentioned in [3]. Smart campus technology within the context of the 5G community is mentioned in [4]. The capacity programs of the AR era in schooling are reviewed in [5]. The proposals in those evaluations are era-pushed, while the principal unit of schooling, students, and teachers is not always targeted on any such clever era campus. Figure 1 suggests the everyday shape of a smart campus with key technology that aids its operation.

The relaxation of the report is prepared as follows. Assistive technology inside the clever campus is mentioned in Sect. 2; the imaginative and prescient of the human-centric, gaining knowledge of-orientated clever campus is contained in Sect. 3; feasible smart services are explored in Sect. 4; the interdisciplinary implications of the clever campus are mentioned in Sect. 5; and Section 6 concludes the report.

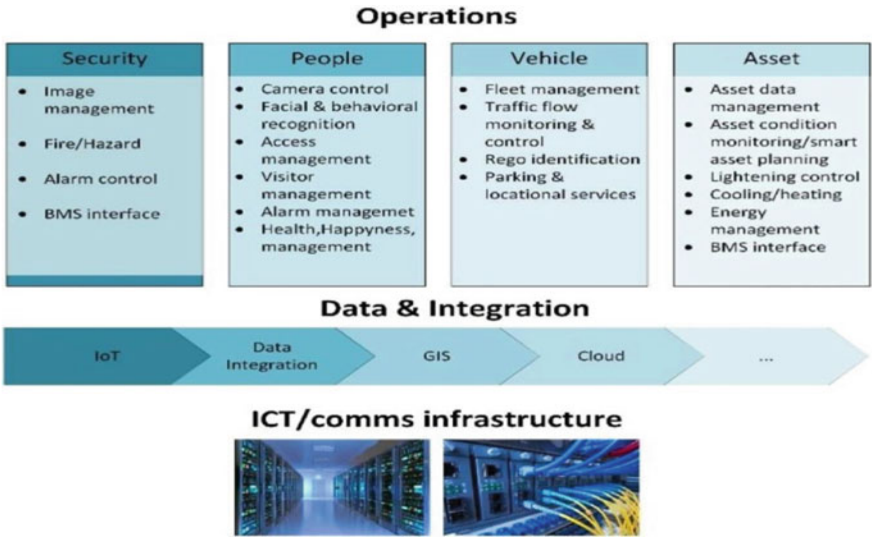


Fig. 1 Smart campus architecture

2 The Technologies of Smart Campus

The improvement of the clever campus might now no longer be feasible with out technological innovations. In the literature, cloud computing, IoT, AR, and AI are some of the most important technology helping the clever campus revolution. The precept of those technology and their blessings for the clever campus is mentioned in this section.

2.1 Cloud Computing

It is a dispensed computing version that gives convenience, the on-call community gets admission to a shared pool of configurable computing sources (which includes networks, storage, and applications) that scales swiftly and may be deployed and deployed. And sharing is primarily based totally on the user. Request with minimum interplay with the provider [6]. The popularization of cloud-primarily-based totally systems has been diagnosed as a key fashion within the discipline of technology-more advantageous clever studying. Compared to conventional computing infrastructure, where each hardware and software program are owned and maintained by businesses on-site, cloud computing permits studying activities in an unstructured environment. It permits inexperienced persons to have short get admission to online studying sources and offerings anytime, anywhere, with limitless scalability, more comfort, and decreased costs [7]. By the usage of a cloud-primarily based total studying platform on the clever campus, digital studying substances may be created and shared seamlessly, increasing the temporal and spatial dimensions of coaching and studying and facilitating studying activities.

2.2 Internet of Things (IoT)

By integrating the user's digital gadgets, clever sensor gadgets, the net, and superior communique technologies, the IoT extends net connectivity to bodily gadgets and ordinary objects. The destiny computing paradigm is expected to transport past the conventional cellular mode primarily based totally on smartphones and laptops and toward an environment surrounded by related and shrewd objects [8].

The capacity blessings of imposing IoT technology in clever campuses especially lie in 3 aspects. First, IoT gives statistics to online educators to track pupils gaining knowledge of development and take knowledgeable action. Second, IOT automates the clever campus operation of the campus and simplifies the learning of the process.

2.3 Augmented Reality (AR)

AR is a brand new shape of revel in wherein the actual international is augmented through digital content material from a computer, permitting for a unbroken overlap and fusion among computer-generated content material and our belief of the actual international [9]. As a next-era interface, AR gives a one of a kind way of interacting and collecting reviews to beautify the coaching/studying surroundings. On an AR-powered smart campus, college students usually advantage a higher understanding and knowledge of what's taking place round them, which complements their studying reviews.

2.4 Artificial Intelligence (AI)

It is a computation technological know-how of creating machines or structures to study from experience, adapt to new inputs, and carry out human-like tasks, which may be a the proper method to resolve troubles in which solutions may be now no longer regularly generated through analytical analysis. Based at the perceived environment, the finished AI set of policies must be capable of maximizing the chance of the agent to successfully attain its cause through interplay with the environment or extracting vital data from statistical data. AI has currently received huge success in lots of actual global applications, which include sample recognition [10], forecasting, translation, control, and games.

3 Human Learning Smart Campus (HLSC)

Smart campuses act as a key vicinity of the smart city, and they may be often placed in a similar socio-economic, environmental, and geographic context, due to this that they percent not unusual place infrastructures, communication channels, similar services, delivery networks, or may be disturbing conditions and dreams. As noted in Sect. 3, the implementation of a smart campus can in an element draw on the experience of various smart city areas, resulting in some smart applications that can be universally required, together with energy management, waste management, health management, sustainability, etc. However, because of the truth, the campus is a place of transport of educational services with university college students and instructors as its cornerstone, it would make greater enjoyable to mix the voice of university college students and instructors into the smart layout of the campus and cognizance of the increase and improvement of university college students and to decorate the excellent of education.

3.1 Design Criteria

Based at the evaluate of current work, we summarize the subsequent standards that must be taken into consideration while enforcing rising technology with inside the clever campus.

3.1.1 Human-Centered

Human-targeted layout is described as an method to device layout and improvement that pursuits to make interactive systems extra beneficial via way of means of specializing in device utilization and person wishes in [12]. This method emphasizes human experience, pride, and performance, which improves the effectiveness and performance of the device in human-associated activities, while neutralizing feasible negative consequences on health, human protection, and performance. From an imperative factor of view, a clever campus is an academic group concerning diverse stakeholders. The important clever campus stakeholders commonly consist of students, instructors, parents, and management teams, all of whom tackle special roles. Depending on their duties and obligations, actors' expectancies of campus intelligence may also be diverted. The human-targeted criterion way that the improvement of the clever campus has to now no longer simplest be student-targeted, however, also do not forget the pursuits of different stakeholders and strive to meet their wishes in a coordinated manner.

3.1.2 Learning-Oriented

The smart campus can be viewed as a simplified model of the smart city, covering several areas including the six pillars of intelligence proposed in [11]. According to the survey sequences in [13], educational institutions choose to invest more in smart learning than other factors such as health, social, energy, management, or governance. Smart campus should be about learning.

There is proof that integrating rising technology right into a clever campus can offer amazing possibilities for college students to research in essentially exceptional methods than conventional instructional. For example, network-based technology permits college students to research independently at domestic and/or through a digital platform, which gives interaction with the actual international and the virtual international with sufficient mastering support. The recognition of cellular gadgets brings amazing comfort to get data anytime, anywhere, which helps ubiquitous mastering. Some studies report that sure activities historically taken into consideration recreational, which include gambling games, social networking, and watching movies at the moment are additionally instructional techniques to manual pupil improvement [14].

3.1.3 Interdisciplinarity

A clever campus is not always an isolated system, however, an critical a part of a clever town. The clever town improvement plan is usually multidimensional, masking the a couple of disciplines that aid the lives of citizens. The broadly ordinary taxonomy of clever town dimensions is given in [15] as clever economy, clever humans, clever governance, clever mobility, clever surroundings, and clever life. Since the improvement and high-satisfactory of clever humans in a town strongly relies upon at the schooling they have got received, the clever campus, as an group supplying instructional services, serves as the idea for the peak of clever humans.

3.2 Definition of Institution

Although the idea of a clever campus becomes discussed many years ago, there has been no regularly occurring and clean definition of it. The improvement of a clever campus will not be centered with out a not unusual place knowledge of what precisely a clever campus is. Existing clever campus proposals are in particular technology driven and miss cross-disciplinary factors. Keeping the layout standards in mind, we envision the destiny campus as HLSC that is described as an academic surroundings penetrated with wise service-allowing technology to enhance academic overall performance with the aid of using serving stakeholder interests, with enormous interactions with different interdisciplinary fields in the clever town context.

3.3 Structure Design

The shape of the HLSC is shown in Fig. 2, where a smart campus plays a crucial role in the context of a smart metropolis to provide academic opportunities for the younger generation. It is also linked to different regions within a smart metropolis, including economy, society, legislation, environment, politics, etc., so interdisciplinary elements can either limit or promote the improvement of a smart campus. This shows the interdisciplinary nature of the intelligence campus. The main framework of a smart campus includes 3 phases surrounding the stakeholders: the infrastructure layer, the technology layer, and the provider layer. The outermost layer serves as the underlying infrastructure of a smart campus, while the innermost layer consists of the factors that can be directly realized and influenced by stakeholders. The framework puts stakeholders at the center, which indicates that sport is practiced at all levels of the smart campus, although a number of them are not directly associated with stakeholders, should be focused at the hobbies of stakeholders. Stakeholder wishes and the 3 tiers of the clever campus are similarly defined as follows:

Fig. 2 Proposed structure for HLSC

3.3.1 Stakeholders of Smart Campuses

The design, construction, safety, and operation of smart campuses comprise the participation and engagement of a pair of stakeholders, which includes students, academic staff, noneducational staff, dad and mom, and manage teams. Therefore, feedback from the ones stakeholders is of outstanding importance to the improvement of the HLSC. In general, due to the awesome roles of stakeholders, their wishes and contributions to campus intelligence differ, so know-how the stakeholders is a must, essential to maximizing the fee of the HLSC.

3.3.2 Infrastructure Layer

A proper supporting infrastructure is important with in the development of smart campuses, due to the fact that it is far described as the muse for the alternative layers. It want to now not simplest embody the ICT elements that assist new technology and are in accordance with the smart concept but moreover contain people as a part of the infrastructure. Essential elements of ICT embody, but are not confined to, records sensing devices, records processing equipment, storage, and harassed and exchange network. People proper right here mainly seek advice from personnel who layout, assemble, and keep the infrastructure and system. As the smart campus is deeply penetrated via way of means of new technologies, handiest humans with technological qualifications are able to control such an infrastructural reinforcement. Without them, campus structures could now no longer be capable of operating as correctly as they need to.

3.3.3 Technology Tier

It represents the middle layer in the smart campus framework, as proven in Fig. 2. Although now no longer without delay associated with mastering, this technology layer is based on infrastructure to create the surroundings wherein sensible mastering occurs and additionally serves as an academic catalyst that allows the transformation of mastering. Conventional mastering into sensible mastering, which overcomes the boundaries of conventional schooling, consisting of time and area constraints and monotonous schooling mode that should prevent man or woman skills and capacity improvement.

3.3.4 Service Tier

Service tier includes smart campus programs that may be carried out immediately to involved parties. In HLSCs, those offerings supplied should be capable of reply to the desires of various actors, with the purpose of enhancing academic performance. The human-centric clever campus idea calls for the carrier company to understand and reply to the intelligence desires of various stakeholders. Knowledge of stakeholder desires might be acquired from an nameless survey of every sort of stakeholder to define use instances to be able to shape a preferred database to guide the improvement of the HLSC. To genuinely replicate biased data approximately the user's case, the survey ought to be designed multidimensional to distinguish schooling levels and geographical data and up to date regularly. In the meantime, ordinary opinions of present clever offerings are also had to tune stakeholder desires.

4 HLSC Services

In HLSC, the supply of clever campus offerings is predicted to without delay or circuitously clever campus student studying outcomes. Learning-orientated clever campus offerings may be divided into 3 classes in keeping with their importance, functionality, and goal users, which might be essential offerings, personalized offerings, and supplementary offerings. The offerings that can be supplied in every class are precise on this section.

4.1 Essential Services

Essential offerings talk over with critical clever capabilities furnished through the campus, commonly carried out to all college students and staff. The capability critical offerings are summarized as follows:

4.1.1 Service of the Physical Environment

The situations of the bodily surroundings in an area will at once have an effect on the consolation, cognition, and fitness of these involved. Physical environment service refers to the real-time calibration of key physical environmental factors, including lighting, temperature, and humidity, to create a comfortable and green environment through the use of IoT technologies. This provider of the bodily surroundings now no longer simplest guarantees the inner learning and residing consolation of the constructing however additionally targets to limit carbon emissions and make a contribution to power performance and environmental sustainability [20]. In addition, the bodily surroundings may be contextually optimized in phrases of the situations of these present, to assist enhance the learning enjoy of students.

4.1.2 Security Service

As a cyber-a-bodily machine, a clever campus calls for safety offerings from each a bodily and IT perspective. Physical security generally refers to the computerized analysis of video from surveillance cameras located in public areas of a campus, the real-time monitoring of moving objects, and the extraction of vital records using intelligent techniques. Once capacity safety risks are assumed, an early caution must be triggered, which fast triggers safety manage measures to be executed with the aid of using safety personnel. By making use of the bodily safety provider, safety incidents might be avoided earlier, and the fake alarm resulting from human intervention is eliminated, which affords a safe and stable bodily surroundings for teaching/gaining knowledge of.

4.1.3 Management Service

In the fully IoT-based infrastructure, personalized information and information about physical aids can accumulate on campus. Based on this multimodal insight, the control provider is designed to toggle the way stakeholders interact with campus assets in three ways. The first is to intelligently allocate spatial resources along with classrooms, offices, meeting rooms, and living spaces. The secnod is to manipulate the deliver and use of strength assets in real-time to fulfill non-public desires and optimize strength savings on campus. The third is time aid control, which refers back to the right making plans of campus sports to optimize the getting to know/paintings performance of stakeholders.

4.1.4 Navigation Service

The smart campus is initially equipped with surveillance cameras for security purposes. Meanwhile, it gives a wealth of records indicating human beings's vicinity over time, imparting the cap potential to perceive specific human beings from video and song their fingerprints the usage of facial reputation technology. Based at the non-public footprint, a campus navigation provider for workforce should offer seamless navigation on and stale campus and quick find wherein events are going on and those who want help [22]. Attempts to offer peace of thoughts provider to human beings on campus so that they could cognizance greater on their studying activities.

4.1.5 AR Service

AR era can offer a unbroken connection among digital content material and the actual international environment. The subsequent query is what sorts of offerings have to AR offer at the clever campus to decorate the coaching/getting to know experience?

In [9], the five sorts of AR-primarily-based totally offerings are identified: AR books, AR video games, AR item modeling, and AR labs. Would be to be had for college kids. An AR ebook at the existence cycle of bugs has been designed and examined on technology college students at an primary faculty in Taiwan, which exams the capacity of the AR-primarily-based totally ebook to encourage creativeness capacity college students and in addition enhance their motivation to learn.

4.1.6 Lab Service

Using an IoT-based lab environment and AI-based lab devices, target labs should be equipped with smart lab devices capable of proactively interacting with college students. One example is that labware can automatically provide real-time academic lab performance feedback and guide college students in completing their lab assignments. This allows college students to benefit from an intuitive experience of human-laptop interaction in the form of audio, video, and AR, so that they can fully concentrate on their lab tasks and achieve a great lab experience. Additionally, digital and remote labs can be a realistic and green strategy to support experiential teaching on subjects imprisoned by a luxurious system and hard time [23]. Innovative lab offerings can significantly improve the performance of your labs and also save you from many capacity protection problems.

4.1.7 Ubiquitous Learning Service

It pulls the maximum famous mastering substances from the ubiquitous learning service. It establishes an excellent connection among students, instructors, specialists, and different instructional companions round the international and paperwork a dynamic mastering circle to collect information quicker and easier. The everywhere majoring supplier creates a 4A (anyone, anytime, anywhere, any device) environment for students [24], absolutely getting rid of limitations to mastering and definitely realizing.

4.2 Personalized Services

Personalized offerings consist of offerings tailor-made to individuals. In personalized offerings, the content material supplied might be special for anybody on campus. Based on current technologies, the subsequent custom designed offerings are planned.

4.2.1 Smart Card Service

The clever card provider affords every fascinated birthday celebration with a card to update numerous college cards (pupil card, constructing get entry to card, library card, health,card, parking card, etc.), which affords personalized efficiency, convenience, and protection for stakeholders and promotes control standardization throughout college departments. The clever card provider operates especially in 4 areas: private identification, monetary control, public information, and intake monitoring [25].

4.2.2 Social Media Service

Providing social media services calls for top records mining technology. Some records mining equipment concentrated on social media were evolved with inside the literature. For example, in [26], an in-intensity subject matter modeling technique is proposed to come across topics from records-pushed multi-modal non-public microblogs containing texts, images, and videos. In [27], a real-time controlling gadget is proposed to screen and examine Weibo public opinion submitted with the aid of using college students on most important events.

4.2.3 Personalized Learning Service

Based on virtual recordings for the duration of scholar gaining knowledge of activities, the smart campus machine can offer a customized idea map for every scholar.The concept map is up to date in actual time to symbolize how college students accumulate new information and synthesize it into their current information that is vital in fixing the trouble of information personalizing. This records might be used to identify precise scholar gaining knowledge of consequences with remarks from instructors. By studying every scholar's customized gaining knowledge of experience, the machine may want to create and replace the scholar information map, verify and expect scholar gaining knowledge of consequences [28], advocate substances and suitable gaining knowledge of resources, and optimize every scholar's gaining knowledge of routine.

4.2.4 Psychological Service

Psychological nation, along with emotion, is any other aspect affecting the teaching/studying overall performance of instructors and college students. The mental nation of people on campus may want to probably be captured via way of means of studying their physiological statistics the use of AI based emotion popularity technology. Preliminary emotion popularity methodologies had been advanced rapidly in current years, along with speech emotion popularity, facial emotion popularity [29], and multi-modal emotion popularity. By making use of emotion popularity recursively in actual timed service psychological nation of instructors and college students might be monitored, which have to offer sufficient proof to provide an explanation for how the mental aspect impacts teaching/studying overall performance at the clever campus.

4.3 Supplementary Services

Supplementary offerings are extra alternatives that would be delivered to critical or personalized offerings, which can also function an awesome interface for enhancing campus offerings and adapting to new technologies. Some examples are supplied as follows:

4.3.1 Reminder Service

The reminder service video display units the calendar of campus and private sports and adaptive sends reminders to customers beforehand of events. It also can offer healthful reminders primarily based totally on people's condition, such as reminding them to rise up and relaxation after long-term sitting or running with inside the lab. The callback provider is typically provided as an alternative that may be became off via way of means of customers if they are sure in their agenda and fitness status.

4.3.2 Extracurricular Activities Service

Participation in extracurricular sports dietary supplements getting to know sports and gives college students with essential possibilities to increase social engagement, find out new interests, and encourage creativity. The extracurricular interest ambitions to evolve the off-Web Website online sports to the special companies of college students. Students are grouped based totally mostly on their availability, background, and preferences, while hobby data collectively with topics, destinations, routes, and packages are designed the use of IoT and social media. During sports, the smart hobby issuer may additionally display screen the hobby repute and dynamically modify plans to govern the activation method to optimize participant pride while ensuring safety and organization of employees safety.

4.3.3 Robotic Service

Robotics as a physical representation of synthetic intelligence is becoming increasingly ubiquitous and likely relevant to many areas of life. In situations where there is no permanent human accomplice robotics are always available and can provide task-based feedback and motivating learning aids to improve student learning experiences [30]. Robotics serves as a supplement to the personalized getting to know provider.

4.4 Service-Actor Interactions

In HLSC, similarly to summarizing the potential shrewd offerings, it is also essential to investigate how the extraordinary offerings are linked with the actors. Our purpose is to reply the subsequent questions:

1. Which stakeholders can gain from the usage of every service?
2. How do stakeholders have interaction with inside the clever campus device?
3. How do offerings enhance the interplay among stakeholders?

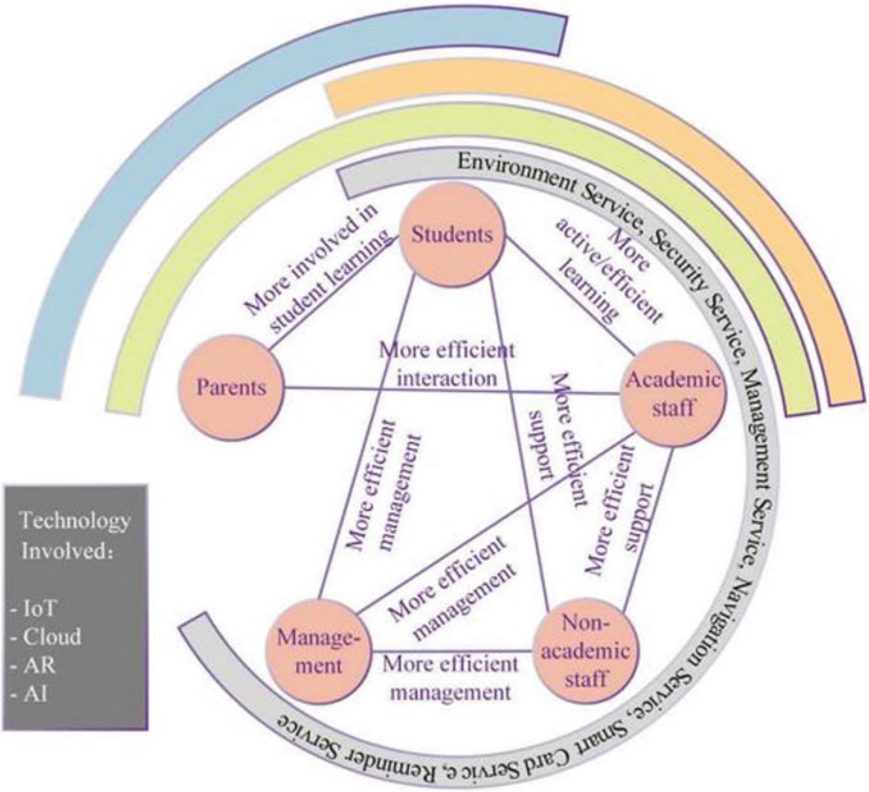


Fig. 3 Interactions between smart campus services

Interactions among stakeholders and clever campus offerings are illustrated in Fig. 3, in which every shape of partial loop covers the stakeholders who may want to gain from the usage of the blanketed offerings. Links among stakeholders indicate how their interactions may be advanced the usage of shrewd offerings. For example, the offerings blanketed with inside the gray shape of the partial ring will be powerful on 4 sorts of stakeholders, specifically students, instructional personnel, non-instructional personnel, and control teams.

5 Conclusion

As a high-give up shape of a clever schooling system, the clever campus is receiving growing studies interest worldwide. Based on a complete evaluate of helping technology and associated clever campus work, this examine predicts, defines, and frames HLSC, that is described as an academic surroundings infused with technology permitting clever offerings to beautify instructional performance while meeting the hobbies of stakeholders, with many interactions with different interdisciplinary fields inside the context of the clever city. Infrastructure, technology, and carrier are diagnosed because the 3 critical layers of the clever campus framework, all of which must be centered on the pursuits of the stakeholders involved. Context-aware, data-driven, forward-looking, immersive, collaborative, and ubiquitous are recognized due to the fact the six center trends of the smart campus. Potential smart services to be provided on campus had been explored and move disciplinary factors that promote or limitation the development of smart campuses had been moreover discussed.

References

1. Liu D, Huang R, Wosinski M (2017) Smart learning in smart cities. Springer, Singapore
2. Global Smart Education Market 2022–2026, 20 Dec 2022. Available at https://www.researchandmarkets.com/reports/4894432/global-smart-education-market-2022-2026pos-0
3. Baldassarre MT et al (2018) Cloud computing for education: a systematic mapping study. IEEE Trans Educ 61(3):234–244
4. Xu X et al (2019) Research on key technologies of smart campus teaching platform based on 5G network. IEEE Access 7:20664–20675
5. Cheng P (2017) A review of using Augmented Reality in Education from 2011 to 2016. Innov Smart Learn 13–18
6. Nayyar A (2019) Handbook of cloud computing: basic to advance research on the concepts and design of cloud computing. BPB Publications
7. Ercan T (2010) Effective use of cloud computing in educational institutions. Proc-Soc Behav Sci 2(2):938–942
8. Aldowah H et al (2017) Internet of Things in higher education: a study on future learning. J Phys: Conf Ser 892(1) (IOP Publishing)
9. Yuen SC-Y, Yaoyuneyong G, Johnson F (2011) Augmented reality: An overview and five directions for AR in education. J Educ Technol Dev Exchange (JETDE) 4(1):11
10. Zhang Y et al (2018) Real-time assessment of fault-induced delayed voltage recovery: a probabilistic self-adaptive data-driven method. IEEE Trans Smart Grid 10(3):2485–2494
11. Ng JWP et al (2010) The intelligent campus (iCampus): end-to-end learning lifecycle of a knowledge ecosystem. In: 2010 Sixth international conference on intelligent environments. IEEE
12. International Organization for Standardization (2010) Ergonomics of human-system interaction: Part 210: human-centred design for interactive systems. ISO
13. Aion N et al (2012) Intelligent campus (iCampus) impact study. In: 2012 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology, vol 3. IEEE
14. Chou Chih-Hsuan, Hwang Chueh-Lung, Ying-Tai Wu (2012) Effect of exercise on physical function, daily living activities, and quality of life in the frail older adults: a meta-analysis. Archives of physical medicine and rehabilitation 93(2):237–244
15. Giffinger R et al (2007) City-ranking of European medium-sized cities. Cent Reg Sci Vienna UT 9(1):1–12
16. Yang AM et al (2018) Situational awareness system in the smart campus. IEEE Access 6:63976–63986
17. Hew KF, Cheung WS (2010) Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: a review of the research. Br J Educ Technol 41(1):33–55
18. Burns M, Pierson E, Reddy S (2014) Working together: how teachers teach and students learn in collaborative learning environments. Int J Instruction 7(1)

19. Yahya S, Ahmad E, Jalil KA (2010) The definition and characteristics of ubiquitous learning: a discussion. Int J Educ Dev Using ICT 6(1)
20. Caţă M (2015) Smart university, a new concept in the Internet of Things. In: 2015 14th RoEduNet international conference-networking in education and research (RoEduNet NER). IEEE
21. Sánchez-Torres B et al (2018) Smart campus: trends in cybersecurity and future development. Revista Facultad de Ingeniería 27(47):104–112
22. Chen L-W et al (2018) Smart campus care and guiding with dedicated video foot printing through Internet of Things technologies. IEEE Access 6: 43956–43966
23. Jara CA et al (2011) Hands-on experiences of undergraduate students in Automatics and Robotics using a virtual and remote laboratory. Comput Educ 57(4): 2451–2461
24. Kelly T (2005) The 4A vision: anytime, anywhere, by anyone and anything. In: Presentation at ITAHK luncheon 8
25. Yang C-H (1999) On the design of campus-wide multi-purpose smart card systems. In: Proceedings IEEE 33rd annual 1999 international Carnahan conference on security technology (Cat. No. 99CH36303). IEEE
26. Peng J et al (2018) Social media based topic modeling for smart campus: a deep topical correlation analysis method. IEEE Access 7:7555–7564
27. Nan F et al (2018) Real-time monitoring of smart campus and construction of Weibo public opinion platform. IEEE Access 6:76502–76515
28. Qu S et al (2018) Predicting achievement of students in smart campus. IEEE Access 6:60264–60273
29. Goldman AI, Sripada CS (2005) Simulationist models of face-based emotion recognition. Cognition 94(3):193–213
30. Lubold N, Walker E, Pon-Barry H (2016) Effects of voice-adaptation and social dialogue on perceptions of a robotic learning companion. In: 2016 11th ACM/IEEE international conference on human-robot interaction (HRI). IEEE

BER Analysis Over a Rayleigh Fading Channel: An Investigation Using the NOMA Scheme

Michael David, Abraham Usman Usman, and Chekwas Ifeanyi Chikezie

Abstract Researches have been carried out in the academia and the industry to examine the error performance of the Non-Orthogonal Several Access (NOMA) schemes since NOMA can serve multiple users concurrently while using the same time and frequency resources. Because its access is not orthogonal, interference between users is a fundamental disadvantage of the NOMA technology. An interference cancellation approach, such as successive interference cancellation (SIC) at the receiver, is typically used to resolve this. Contrarily, inter-user interference in the SIC process cannot be eliminated and is usually due to wrong decisions at the receiver caused by the channel. The performance of the downlink NOMA for the BPSK transmission system in a Rayleigh fading was assessed in this paper using MATLAB. The findings demonstrate that NOMA offers users reasonable fairness while minimizing interference at a reasonable BER.

Keywords NOMA · SIC · BER · BPSK

1 Introduction

The fundamental idea of NOMA is to utilize the power domain for multiple access in contrast to previous generations of mobile networks, which depend on the time/frequency/code domain [1]. The fundamental drawback of orthogonal multiple access (OMA) approaches is that they have a low spectral efficiency when some bandwidth resources, like subcarrier channels, are given to users with low channel

---

M. David (✉) · A. Usman Usman · C. I. Chikezie
Department of Telecommunications Engineering, Federal University of Technology, Minna, Niger State, Nigeria
e-mail: mikeforheaven@futminna.edu.ng

A. Usman Usman
e-mail: usman.abraham@futminna.edu.ng

C. I. Chikezie
e-mail: chekwas.pg2010285@st.futminna.edu.ng

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_84

state information (CSI). However, while employing NOMA, every user has access to every subcarrier channel. Thus, the bandwidth resources allotted to users with low CSI can still be accessed by users with high CSI, thus increasing spectral efficiency [2]. Superposition coding at the transmitter and Successive Interference Cancellation (SIC) at the receivers are the key components of NOMA, which is anticipated to outperform Orthogonal Multiple Access (OMA) in terms of spectral efficiency [3].

For optimal performance, signal transmission attenuation, distortion, and noise must be minimized. The transmitting and receiving signals must therefore be accurately measured. Factors, coding, features, and various digital modulation techniques can impact the reliability of the received signal and the transmission quality. In contrast to its wired counterpart, wireless technology has several advantages, such as enhanced mobility, higher productivity, reduced costs, simpler installation, and scalability [4]. As a result of reflection, diffraction, and scattering effects, transmitted signals arrive at the receiver with varying power and delay, which is one of the limitations and drawbacks of different transmission channels in the wireless medium between the transmitter and receiver.

When data is transmitted over a wireless channel, there is a risk of errors in the system. The system's integrity can be compromised if errors are introduced into the data [5]. Therefore, evaluating the system's performance is necessary, and the bit error rate BER provides an ideal method to achieve this goal. Unlike many other types of evaluation, BER evaluates the end-to-end performance of a system, including the transmitter, receiver, and mediation between the two. In this way, the BER can test the system's actual performance rather than testing the components and hoping they perform satisfactorily once they are in place [6]. The BER value for the wireless medium is relatively high. The efficiency of wireless data transfer may suffer from these problems. Error management is therefore required for many applications.

Using discrete signals, a carrier wave is modified using the digital modulation approach. High carrier frequencies are employed in digital modulation to facilitate signal transmission over long distances using existing long-distance communication methods, such as radio channels [7]. The received demodulated signal is not adversely affected by channel noise. Conversely, the demodulated signal is distorted if the analogue signal contains noise. Applications that run on the fifth generation (5G) radio access networks have extremely high speeds, low latency, mass connectivity, and good mobility. [8, 9]. NOMA enables high-density networks and great spectral efficiency by allowing users to access the same radio resources [10]. Multiple users are served by conventional Orthogonal Multiple Access (OMA) schemes by assigning them to various radio resources, such as frequency and time. Unlike OMA, which splits users into power domains, NOMA services large numbers of User Equipment (UE) concurrently on the same resource blocks. Superposition coding at the transmitter and successive interference cancellation at the receiver are the fundamentals of a NOMA technique [11, 12]. Figure 1 details the operation of a digital communication networks.

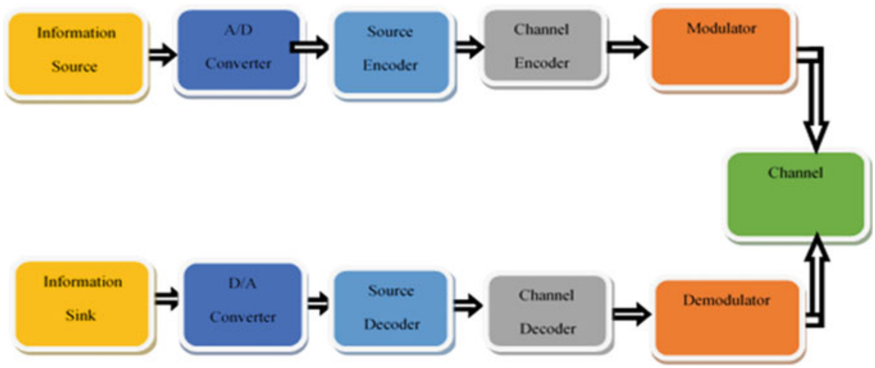


Fig. 1 A digital communication system's block diagram

2 System Model

A wireless channel is vulnerable to fading and multipath propagation. Numerous channel models can be used to capture the effects of fading. Every model aims at a specific circumstance. The Rayleigh fading model is one example. The Rayleigh fading model can be used when there is no line of sight (LOS) path between the transmitter and the receiver. As a result of reflection, scattering, diffraction, and shadowing, all multipath components undergo small-scale fading. In an extreme form of Rayleigh fading, caused by multipath transmission, every bit transmitted experience a different attenuation and phase shift. In other words, the channel changes for every bit. The Rayleigh fading model is used to statistically analyse radio signal propagation. It works best without a dominant signal, which often happens with cell phones used in dense urban environments. Figure 2 depicts the network model for a Rayleigh fading channel.

The weak user in NOMA is given additional transmission power. By interpreting the messages of other users as noise, the weak user can decode its message [2]. On the other hand, the strong user will first identify its message partner under the stronger channel state, subtract the message from the weak user, and last decode its own message. This method explains the successive interference cancellation.

The base station has two discrete messages x_f to the far user, and x_n to the near user. The power allocation factors are α_f and α_n respectively, for the far and the near user (where $\alpha_f + \alpha_n = 1$). In a NOMA system, more power is allocated to the far user and less to the near user to promote user fairness ($\alpha_f > \alpha_n$).

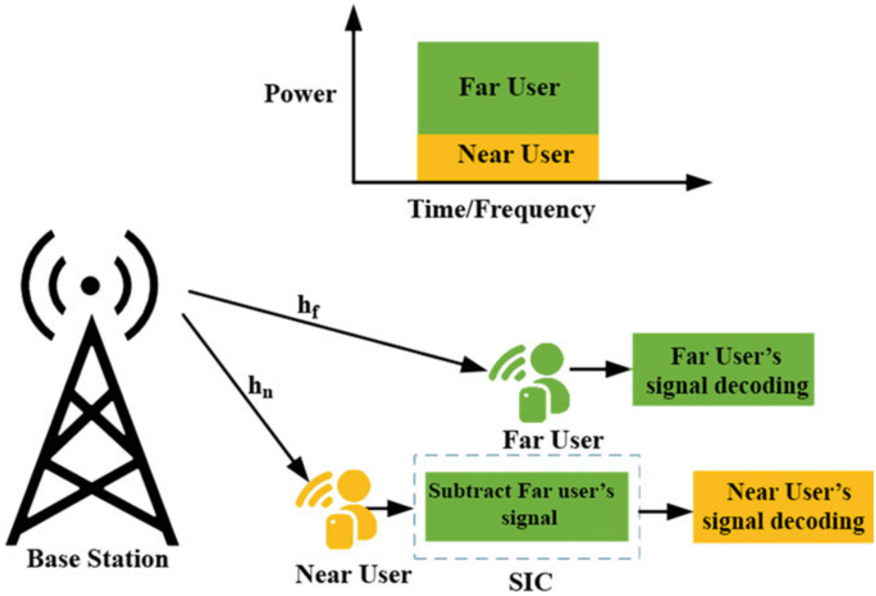


Fig. 2 Network model

2.1 NOMA Encoding and Transmission

The base station transmits a superposition-coded NOMA signal that is:

$$x = \sqrt{P}\left(\sqrt{a_f x_f} + \sqrt{a_n x_n}\right) \tag{1}$$

where P is the transmit power.

After propagating through the channel h_f, the copy of x that the near user receives is given as:

$$y_f = h_f w + w_f \tag{2}$$

where w is noise.

Similarly, the copy of x that was propagated through h_n and received by the far user is given as:

$$y_n = h_n w + w_n \tag{3}$$

2.2 NOMA Decoding at the Far User

Expanding the signal received by the far user:

$$y_f = h_f x + w_f \tag{4}$$

$$= h_f\sqrt{P}\left(\sqrt{\alpha_f x_f} + \sqrt{\alpha_n x_n}\right) + w_f \tag{5}$$

$$= h_f\sqrt{P}\left(\sqrt{\alpha_f x_f} + h_f\sqrt{P}\sqrt{\alpha_n x_n}\right) + w_f \tag{6}$$

where

$h_f\sqrt{P}\sqrt{\alpha_f x_f}$ is the desired and dominating signal,

$h_f\sqrt{P}\sqrt{\alpha_n x_n}$ is the interference and low power signal,

w_f is noise.

Direct decoding of y_f would yield x_f since $\alpha_f > \alpha_n$. The term x_n component was considered as an interference. For far user, the signal-to-interference noise ratio is given as

$$\gamma_f = \frac{|h_f|^2 P\alpha_f}{|h_f|^2 P\alpha_n + \sigma^2} \tag{7}$$

and its achievable data rate is given as:

$$R_f = \log_2\left(1 + \gamma_f\right) = \log_2\left(1 + \frac{|h_f|^2 P\alpha_f}{|h_f|^2 P\alpha_n + \sigma^2}\right) \tag{8}$$

2.3 NOMA Decoding at the Near User

Expanding the signal received by the near user:

$$y_n = h_{fn} x + w_n \tag{9}$$

$$= h_n\sqrt{P}\left(\sqrt{\alpha_f x_f} + \sqrt{\alpha_n x_n}\right) + w_n \tag{10}$$

$$= h_n\sqrt{P}\left(\sqrt{\alpha_f x_f} + h_n\sqrt{P}\sqrt{\alpha_n x_n}\right) + w_n \tag{11}$$

where

$h_n\sqrt{P}\sqrt{\alpha_f x_f}$ is the desired and dominating signal,

$h_n\sqrt{P}\sqrt{\alpha_n x_n}$ is the interference and low power signal,
w_n is noise.

Before decoding his own signal, the near user must first perform successive interference cancellation (SIC). The SIC procedures are as follows;

1. direct decoding of y_n obtains x_f or more specially, an estimate of x_f, which is $\bar{x}$
2. $y_n' = y_n - \sqrt{\alpha_f \overline{x}_f}$ is computed
3. y_n' is decoded to obtain an estimate of x_n

Before SIC, the signal-to-interference noise ratio at the near user for decoding the signal of the far user is given as

$$\gamma_{f,n} = \frac{|h_n|^2 P\alpha_f}{|h_n|^2 P\alpha_n + \sigma^2} \tag{12}$$

The corresponding achievable data rate is given as follows:

$$R_{f,n} = \log_2\big(1 + \gamma_{f,n}\big) = \log_2\left(1 + \frac{|h_n|^2 P\alpha_f}{|h_n|^2 P\alpha_n + \sigma^2}\right) \tag{13}$$

3 BER of a NOMA System

Firstly, we declared the values of some parameters. For the distances, $D_f = 1000\,\text{m}$, and $D_n = 500\,\text{m}$. Then we set the power allocation factors as $\alpha_f = 70$ and $\alpha_n = 30$. For user fairness, we allocated more power to the far user. We initialized a range of 0–40 dBm for the transmit power. Our system's bandwidth was then set to $B = 1\,MH_z$. According to the formulae, $N_0 = kTB$, where $k = 1.38 \times 10^{-23}$(Boltzmann constant), $T = 300K$, the thermal noise power was calculated. We then generated the Rayleigh fading coefficients for h_f and h_n. We set the path loss exponent $\eta = 4$. Next, we generated noise samples for the far and near users and randomized binary data for the users. We calculated the superposition-coded signal x after using BPSK to modulate the data. We also calculated y_f and y_n and then equalized them by diving h_f and h_n respectively. From the equalized version of y_f, we performed direct BPSK demodulation to obtain $\overline{x}_f$. We used the biterrr function to estimate BER and compared x̄ with the original data from the far user. To estimate x_f, we directly decoded the equalized version of y_n. We decoded the signal to obtain, $\overline{x}_n$ by remodulating x_f and subtrahend the remodulated x_f element from the equalized version of y_n. We further compared $\overline{x}_n$ with the near user's initial data, we estimated BER using the biterrr function. Finally, we plotted the BERs in relation to transmit power using MATLAB.

The BER performance for a two-user scenario is shown in Fig. 3. The near and far users were allocated 0.70 and 0.30 power respectively, with a 1 MHz bandwidth using the BPSK modulation technique. According to the figure above, interference

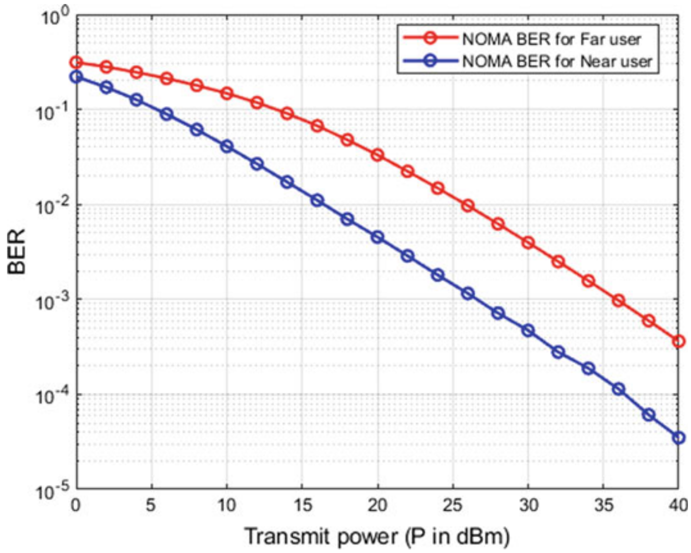


Fig. 3 Theoretical and simulated BER performance

Table 1 NOMA BER analysis

Transmit power (dBm)	Far user	Near user
10	0.14857	0.040852
20	0. 033698	0.00449
30	0.004079	0.000422
40	0.000447	0.000037

from the near user causes the far user to have a greater BER. With no interference, the near user has the lowest BER. This shows that NOMA performs as expected as shown in Table 1 below.

4 Conclusion

The integrity of the information transmitted through the downlink NOMA system can be assessed using the BER of a digital signal, which is a crucial metric. This work used MATLAB to evaluate the BER performance of a downlink NOMA with a BPSK transmission scheme over a Rayleigh fading channel. The result demonstrated that NOMA offers users a fair system that is acceptable while minimizing interference and maintaining a reasonable BER.

References

1. Chikezie CI, David M, Usman AU (2022) Power allocation optimization in NOMA system for user fairness in 5G networks. In: Proceedings 2022 IEEE Niger 4th international conference disruptive technology sustainable development. NIGERCON 2022. https://doi.org/10.1109/NIGERCON54645.2022.9803107
2. Ding Z et al (2017) Application of non-orthogonal multiple access in LTE and 5G networks. IEEE Commun Mag 55(2):185–191. https://doi.org/10.1109/MCOM.2017.1500657CM
3. Saito Y, Benjebbour A, Kishiyama Y, Nakamura T (2013) System-level performance evaluation of downlink non-orthogonal multiple access (NOMA). https://doi.org/10.1109/PIMRC.2013.6666209
4. Attaran M (2021) The impact of 5G on the evolution of intelligent automation and industry digitization. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-020-02521-x
5. Hilario-Tacuri A, Maldonado J, Revollo M, Chambi H (2021) Bit error rate analysis of NOMA-OFDM in 5G systems with non-linear HPA with memory. IEEE Access 9. https://doi.org/10.1109/ACCESS.2021.3087536
6. Vasuki A, Ponnusamy V (2022) Error rate analysis of intelligent reflecting surfaces aided non-orthogonal multiple access system. Intell Autom Soft Comput 33(1). https://doi.org/10.32604/iasc.2022.022586
7. Bala D, Waliullah GM, Islam N, Abdullah I, Hossain MA (2021) Analysis of the probability of bit error performance on different digital modulation techniques over AWGN channel using MATLAB. J Electr Eng Electron Control Comput Sci 7(25)
8. Iradier E, Fadda M, Murroni M, Scopelliti P, Araniti G, Montalban J (2022) Nonorthogonal multiple access and subgrouping for improved resource allocation in multicast 5G NR. IEEE Open J Commun Soc 3. https://doi.org/10.1109/OJCOMS.2022.3161312
9. Liu Y et al. (2022) Evolution of NOMA toward next generation multiple access (NGMA) for 6G. IEEE J Sel Areas Commun 40(4). https://doi.org/10.1109/JSAC.2022.3145234
10. Li C, Hu G, Fan B (2022) System-level performance simulation analysis of non-orthogonal multiple access technology in 5G mobile communication network. Int J Commun Syst 35(5). https://doi.org/10.1002/dac.4572
11. Azam I, Shin SY (2022) On the performance of SIC-free spatial modulation aided uplink NOMA under imperfect CSI. ICT Express. https://doi.org/10.1016/j.icte.2021.12.005
12. Hamza AA, Dayoub I, Alouani I, Amrouche A (2022) On the error rate performance of full-duplex cooperative NOMA in wireless networks. IEEE Trans Commun 70(3). https://doi.org/10.1109/TCOMM.2021.3138079

Artificial Intelligent, Digital Democracy and Islamic Party in Indonesian Election 2024

Zuly Qodir

Abstract This article explains that digital democracy, which is now popular in the world of politics, is still too difficult to practice in Indonesia, especially in areas where internet coverage is still difficult. The view that digital democracy will facilitate the political process and increase citizen participation in presidential-vice-presidential elections, regional head elections, and regional representative councils does not seem to be what political policymakers envision. Digital democracy has an impact on the existence of a "democratic elite" who understands online media or the internet. Meanwhile, citizens who are not familiar with the internet will become "democratic Sudras" because they fail to utilize online media. In the upcoming 2024 election, there will be a fight over the use of the internet to spread political ideology and attract citizen's votes. All "democratic elites" and "democratic Sudras" will be associated with Islamic parties in using social media as a campaign for their ideology and political actors.

Keywords Artificial ıntelligent · Digital democracy · Islamic party · Elite democracy · Sudra's democracy

1 Introduction

You The use of the internet and social media is an undeniable trend in the practice of electoral democracy in Indonesia. This kind of tendency, in politics, became known as digital democracy. One of the cores of digital democracy is to use of machines as a voting method in electoral politics [1]. Indonesia as a country with a population of more than 274 million people certainly thinks about whether the practice of electoral politics needs to use machine tools so that the political process can run quickly and with quality, or whether to stick to the conventional political process at an

Z. Qodir (✉)
Department of Islamic Politic – Political Science, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: zuliqodir@umy.ac.id

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_85

expensive cost [2]. The General Election Commission (KPU) also worked for months in conducting the process of counting voter's votes to determine the winner in the presidential election and the legislative candidate who got the most votes.

As is well known to the public, Indonesian democracy is said to be a country with expensive political costs, so political practices become transactional which is manifested in the "buying and selling of voter's votes" in every electoral event [1]. Electoral events from the sub-district, district/city, provincial and national levels are a vehicle for spreading large funds to voters for support. This is a form of democracy full of money politics and kinship politics during the elections after the 1998 political reforms [3]. Citizens participating in political contestations, such as candidates for regional representatives, people's representative councils, village head candidates, regional head candidates, and presidential-vice presidential campaign teams often "take advantage" of citizens with political transactions that run very brightly in each general election after the 1998 reforms [4]. In fact, from 2009 to 2020 the political process became increasingly clear regarding the occurrence of money political transactions. This is where it will be seen how the quality of democracy and citizens are involved in electoral politics. The professionalism of party activists is required to manage the party as well as possible with the condition of money politics that is very dominating [5].

Citizens involved in electoral politics in the digital democracy era will be seen in the quality of understanding the means of democracy, utilizing, and analyzing political events that occur. Acumen paying attention to changes in the political system and the means used will help determine the successful implementation of the electoral democratic process [6]. In addition, digital democracy can be shown to the public regarding freedom of expression, citizen's rights, and political autonomy. Such things are at the forefront of an era called digital democracy [7]. As long as the practice of electoral politics runs above money and transactional politics, the quality of Indonesia's digital democracy can be said to be low. Its benefits are thus insignificant to the development of democracy and the autonomy of citizenship in politics [8].

However, until now, the concept of digital democracy still does not have the agreement of socio-political scientists. Even though the impact of digital democracy has occurred in the political environment, social life, economy, religion, and the autonomy of citizens in determining their suffrage [9]. The issue of citizenship then became a major issue in digital democracy, especially related to the suffrage of citizens and the intervention of machines in determining their political rights. Various political activities of citizens will be seen in the calculation of the use of electronic machines so that the mechanism for counting votes also uses electronic means [10].

If you look at the historical perspective of digital democracy, it can be shown that digital democracy some experts are said to be a "democratic test project", namely from conventional democracy with participation and manual calculations, moving to use electronic machinery [11]. The results can immediately be known who the winner and loser is, but there are difficulties when it comes to knowing the level of fraud and errors in voting, as well as calculations. This is what causes some digital democracy experts to declare it an "ambitious project" of digital age democracy [12]

Digital democracy can be said to be a political activity that slightly ignores the local goodness and wisdom of a country and its citizens. This is because digital democracy is a political activity that uses satellite aircraft channels or channels, be it telegrams, telecommunications, or social media [13]. Where such things have never been known in the practice of electoral politics in the world, let alone in Indonesia. How to use a channel as a substitute for conducting in-person elections by coming to the polling station by only using electronic machine devices to make the choice [14]. Elections can also be conducted anywhere, voting rights holders want to use not have to go to the polling station. This is very different from the electoral practices that have been going on for many years in Indonesia, a country that does not all have adequate channels or internet networks [15].

This article wants to analyze the historical background of the emergence of digital democracy, its socio-political impact on citizens in electoral political participation, the rivalries that occur between parties involved in electoral politics with various ideological backgrounds, as well as the challenges of digital democracy to Islamic parties in Indonesia in the upcoming 2024 elections. This article is based on bibliography data contained in reputable international journals indexed by Scopus from 2009–2022 which were taken using the Nvivo 12-plus program and then analyzed based on the thematic structure discussed in articles related to digital democracy.

2 Method

This study employs big data collected through bibliometrics, which were analyzed using Nvivo-12 Plus. Vos-Viewer was used to visualize the results. Data were collected from the literature, including books, book chapters, research reports, book volumes, and journal articles written by experts/observers interested in digital democracy.

Data was collected from articles in journals, books, and bibliographies that discuss digital democracy. The data was taken by tracing several Scopus database sources, Google Scholar, and connected papers related to digital democracy. This approach was previously pioneered by Tinnes (2021). These questions will be answered by reviewing the topics, frameworks, and previous findings from articles indexed on Scopus. Data collection was carried out through the process of (1) searching for articles, (2) mapping of study topics, (3) analysis of study topics, and (4) conceptualization of digital democracy (Fig. 1).

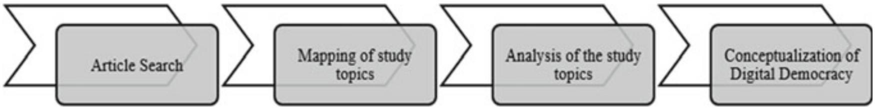


Fig. 1 Research data analyses

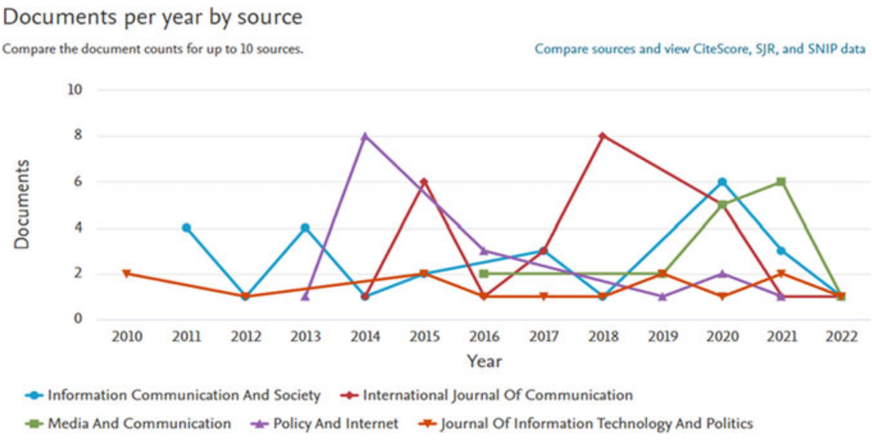


Fig. 2 Journals containing articles dealing with digital democracy

Analysis was conducted in several stages. First, to identify articles, a database of Scopus-indexed articles was searched using the keyword “digital democracy”; approximately 115 articles published between 2010 and 2022 were returned. Second, the researchers verified the relevance and country of origin of the articles, then added them to an Excel table. The articles listed included both widely and rarely cited ones (as measured by H-index). The authors ultimately produced a list of 100 articles that were deemed highly relevant to the research topic. The full texts of these articles were subsequently downloaded, then added to a database together with their year and journal of publication. The results are provided in Fig. 2.

During the mapping stage, the following process was used. First, the full texts of the articles were imported to VOSviewer, a software tool for visualizing bibliometric networks, to identify data clusters and visualize the links between research themes. Second, data were analyzed and conceptualized by reviewing the articles, thereby obtaining the data necessary to answer questions. During this stage, the analysis focused on data clusters, dominant topics, thematic linkages, intellectuals involved, and topic maps. Finally, the researchers examined the reviewed article’s understanding of digital democracy and relevant topics. Figure 3 presents the data collection and analysis process used in this study.

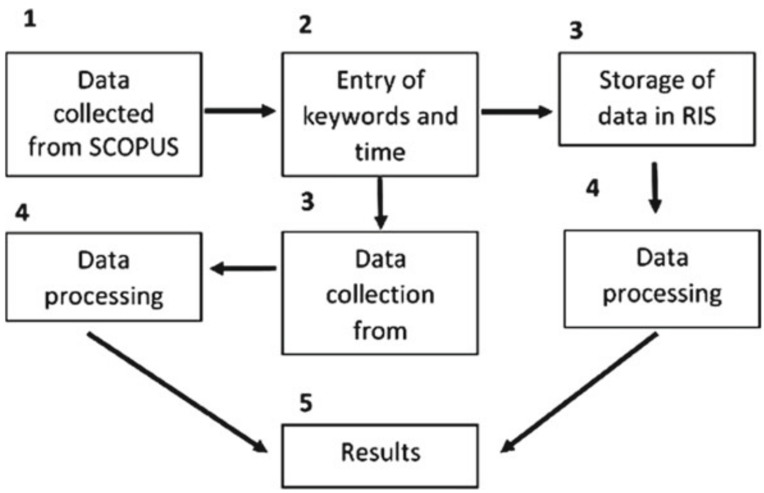


Fig. 3 Data collection and processing

3 Results and Discussions

3.1 Digital Democracy: Difficult Practices Democracy

Words Digital democracy is one of the methods of using tools (machines) or the internet in the practice of organizing democracy. The use of the internet or digital means has started since the 2014 and 2019 Presidential Elections so that the counting of voter's votes runs faster and can be witnessed directly by citizens. This is a form of digital praxis of democracy which is currently one of the objectives of the government in organizing the electoral democratic process [16]. The use of digital is certainly one of the democratic practices that run in various countries, including in developed countries such as America, Australia, Britain, and France.

Apart from the digital democratic process that has been held since the 2014 and 2019 elections in presidential elections, the house of representatives and the regional representative council is safe and fast, it is alleged that the practice of digital democracy has difficulties faced by voters in Indonesia. The difficulties of machine utilization in the practice of digital democracy will be connected to the entire life of citizens including in tourism, community development, and progress and decline of governance in practicing the electoral democratic process. Everything becomes an important issue in a digital democracy [16]. Digital democracy, therefore, has a positive dimension, namely accelerating the voting and calculation process, but the negative dimension is towards areas that do not have a good internet network, as well as changes in community culture from an agrarian society to a machine (electronic) society [17].

Related to Digital democracy in political practices by using the internet, the thing that cannot be forgotten is to strengthen citizens' understanding of the use of the internet or electronic means when conducting elections. Strengthening citizens understanding and master the internet and electronic means is very important as a form of citizen participation in electoral politics [18]. Online media is an option in running electoral politics as a substitute for conventional political practices that have been going on for years in various countries, including Indonesia. This change in the political practice of using electronic machines is also a demand for people to make cultural changes in their lives [6].

In summary, it can be said that electoral politics using digital or internet means is part of strengthening citizens to understand and master technology. Expansion of the use of the internet or online media means of communication to strengthen citizens in voting [19]. Citizen participation in democratic political projects is connected to political ethics, regulations, and laws prevailing in a country and other capabilities in utilizing technology.

The digitization of everyday life enables surveillance from consumer digital devices. Autonomous policing based on sophisticated AI, robotics, facial recognition, and autonomous decision-making may be used for surveillance and/or crowd control. This is where digital democracy in the 2024 electoral election becomes one of the means of succeeding in the political process in Indonesia [20].

If connected with the development of digital democracy that has begun to rise today, we can pay attention to articles that discuss digital democracy in contemporary political practice in the world and Indonesia from 2009–2022 as below. From there we will find that the political development of the digital age will continue to be the mainstream politics of the contemporary era (Fig. 4).

Based on the picture above, can be seen from the development of the study on digital democracy from 2010–2022. The following authors map by year, theme about digital democracy (Table 1).

3.2 Citizen Participation in Election

There is an important issue in the practice of digital democracy related to how to empower citizens to understand and be able to practice digital democracy. Therefore, in the practice of digital democracy, facilities are needed to accelerate the democratic process so that it can run well. Processes that require electronic means can facilitate the election of the president and vice president, the house of representatives, and the regional representative council so that it cannot be abandoned [21]. Even in the case of regional head elections in Indonesia, it must be a concern of digital democracy, so that representative democracy can run well as a regional decision-making option [22]

numbers Digital democracy as a mechanism for citizen participation in elections is a concern for the rights of citizens as a democratic principle in a macro sense. Democratic participation in digital democracy seeks to encourage citizens to be

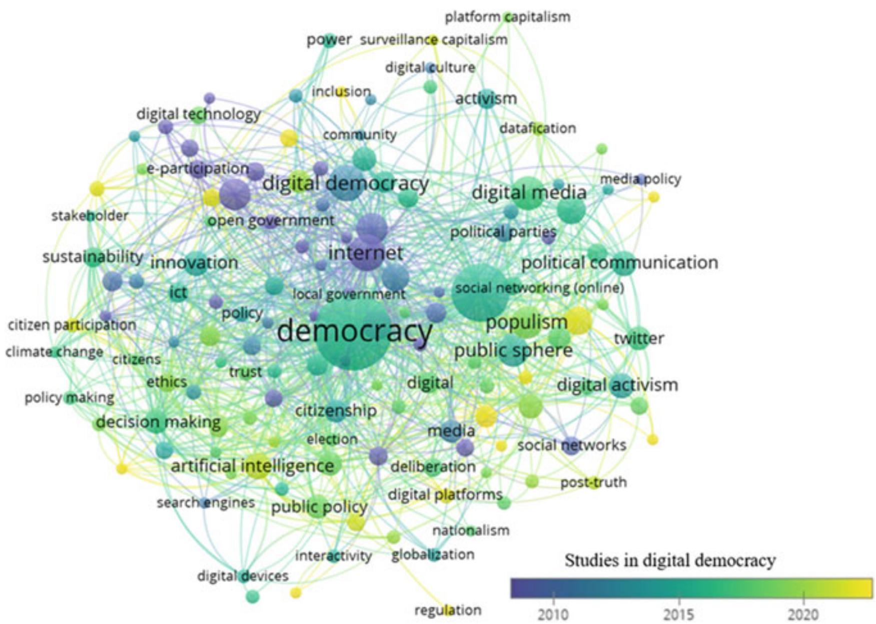


Fig. 4 Development of the study of digital democracy

Table 1 Progress of studies by year, theme, and author

Year	Themes
2010–2016	Globalization, digital democracy, activism, digital culture, political parties, political relations, political system, collection action, communication, civic engagement, social network, digital Inclusion, e-Government, digital divide, new media
2017–2022	Democracy, social media, digital media, digital democracy, innovation, climate change, sustainability, populism, artificial intelligence, disinformation

able to exercise control over public policies, utilize dialogue spaces to criticize the government, evaluate national and local leaders, have sensitivity to the government, and the government is obliged to provide information to citizens about things that have been done and will be done next. [21]. Thus, digital democracy as one of the vehicles for preparing citizens to express their opinions widely in the public sphere becomes visible.

In summary in contemporary political studies, participation in a digital democracy is a conceptualization that wants to practice citizenship participation, representation in the practice of elections and regional elections, and matters related to the limitation of the power of leaders in a country [23]. It can be said in principle that citizen participation in digital democracy is a sustainable democratic culture. The debate over participation in digital democracy will increasingly seriously discuss the rights of citizens, reflect on the journey of democracy critically and provide tools for the running of the electoral process [24].

In such a context, the service to citizens is to be able to utilize technological means, manage data, and consider the possibility of networks using electronic means in conducting general elections and regional head elections [25]. Today the idea of digital democracy by utilizing machine means cannot ignore the culture of a very diverse society.

3.3 Elite and Sudra's Democracy: Representative Democracy

There The most visible problem of digital democracy there will appear groups of people who fall into the category of "elite and sudra" and democracy. This is due to the existence of citizens who do not understand and can use electronic means (machines) and citizens who are already smart to use electronic means. The existence of democratic elites and democratic sudra if not managed properly can create authoritarian politics, which tends to pay less attention to democratic processes and practices. Such things have happened in China, Russia, Turkey, and North Korea [25]

In digital politics, using the internet as a means will lead to the experience of citizens and politicians involved in political parties, companies that help politicians in elections, and to the regions, academia, and civil society are the drivers of a more deliberative and interconnected democracy with various parties, which encourages the democratic process. The existence of internet facilities or technology will cause changes in the election [26]. If all stakeholders carry out their functions properly, the existence of a "democracy gap" that gives rise to elite democracy and sudra democracy will be resolved. However, if stakeholder's democracy does not go well, the possibility of "elite democracy" and "sudra democracy" groups will continue.

In Indonesia, the possibility of two democratic groups needs to be developed innovatively so that the democratic gap can be reduced. Massive strengthening of the public to master and understand technological means such as the internet and social media in the political process needs to continue. The behavior of politicians and the public will increasingly use internet media according to political platforms and the behavior of people who are increasingly active in being very important [27]. Another thing that conventionally needs attention in government is about understanding the importance of democratic practices and democratic values so that their impact on political policy brings the "distance between the political elite" and the "sudra of democracy" closer together [28]. If everything can go well, attention to democratic practices in Indonesia will go well which has legitimacy and political practice.

Digital democracy has an impact on the existence of a "democratic elite" who understands online media or the internet. Meanwhile, citizens who are not familiar with the internet will become "democratic Sudras" because they fail to utilize online media. In the upcoming 2024 election, there will be a fight over the use of the internet to spread political ideology and attract citizen's votes.

3.4 Islamic Party Challengers

The big change from the conventional era of democracy to digital democracy can be said that this is the most real challenge. This suggests an era known as “information disruption” [20]. Public communication that does not use electronic means properly, will be the most real threat from Islamic parties. Islamic parties can be left behind in many political activities such as campaigning and spreading the ideology they carry.

The challenge is particularly evident in the Islamic-based parties that are still conventional in campaigning for programs and promoting party leaders and candidates to be put forward in legislative elections. Therefore, Islamic parties if are not serious about adapting to using electronic means, the delay in disseminating information and designing programs that are following public tendencies, especially young people, will be abandoned by young voters [29]. All uses of electronic media or machine tools in the political process are categorized as the development of Artificial Intelligence Democracy.

Artificial Intelligence Democracy in Indonesian political practice will certainly be a threat to Islamic and nationalist parties that do not pay attention to the development of telecommunications media in politics. If the Islamic party does not pay attention to the development of telecommunications facilities, such as the internet, telegram, social media, and Instagram [30]. Therefore, artificial intelligence can be one of the perspectives in changing the behavior of citizens, political elites, policymakers, and political party leaders.

Overall the practice of digital democracy will give rise to so-called “democratic elites” and “democratic sudra’s”, which will be associated with Islamic parties in using social media as a campaign for their ideology and political actors [31]. This is a serious problem in digital democracy in the upcoming 2024 elections. Some developments in the study of digital utilization by the Islamic parties PKS, PPP, PKB, and PAN in Indonesia can be seen in the chart below (Fig. 5).

The weakness of the Islamic party is not using social media as a campaign tool so that the vision, mission and ideology are shared. Therefore, it is slow to be accepted in society and its spread is not widespread. In the future, Islamic parties will be better able to make maximum use of electronic means to spread their vision, mission and ideology campaigns to the wider community.

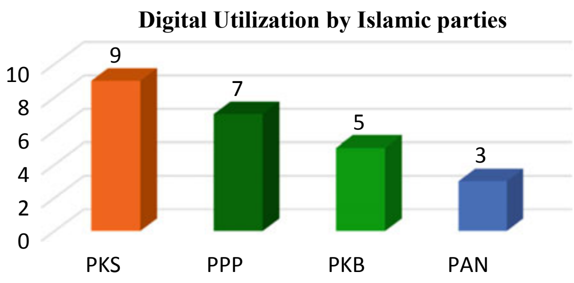


Fig. 5 Digital utilization by Islamic parties

4 Conclusion

By paying attention to the studies in this article, it can be said that artificial intelligence is one of the contemporary forms of utilizing information technology facilities in political practice in Indonesia and even in the world. Utilizing machine (electronic) means in elections can give birth to a category of citizens who are said to be electronic understanding in the "democratic elite" group and the "sudra democracy" community group because they do not understand and are less able to use electronic means.

The practice of democratic digital politics can speed up the process of voting and counting votes, but in practice on the ground, there are the most obvious challenges to some regions or regions where electronic networks and the internet are weak. Therefore, the use of technology in democracy as an era of "digital democracy" can be said to be an era of continued democracy from conventional democracy that utilizes communication technology.

References

1. Mariano-Florentino Cuéllar AZH (2020) Toward the democratic regulation of AI systems: a prolegomenon. 271:1–21
2. Moens P (2022) Professional activists? Party activism among political staffers in parliamentary democracies. Party Polit 28(5):903–915. https://doi.org/10.1177/13540688211027317
3. Aspinall E, Mietzner M (2014) Indonesian politics in 2014: democracy's close call. Bull Indonesia Econ Stud 50(3):347–369. https://doi.org/10.1080/00074918.2014.980375
4. Sales P (2021) Algorithms, artificial intelligence, and the lan. Judicature 105(1):23–35. https://doi.org/10.1080/10854681.2020.1732737
5. Hadiz VR (2013) The rise of capital and the necessity of political economy. J Contemp Asia 43(2):208–225. https://doi.org/10.1080/00472336.2012.757433
6. Posteraro L (2021) The digital-democracy-development nexus: how to effectively advance the EU's s digital policy abroad. Swiss
7. Donahoe E, Metzger MM (2019) Artificial intelligence and human rights. J Democr 30(2):115–126. https://doi.org/10.1353/jod.2019.0029
8. Zimmer B (2018) Democracy under threat: risks and solutions in the era of disinformation and data monopoly. no. December, pp 1–100. [Online]. Available: www.ourcommons.ca
9. Katyal SK (2022) Democracy and distrust in an era of artificial intelligence. Daedalus 151(2):322–334. https://doi.org/10.1162/DAED_a_01919
10. Bundy A (2017) Edinburgh research explorer review of preparing for the future of artificial intelligence Citation for. Ai Soc 32(2):285–287. [Online]. Available: https://doi.org/10.1007/s00146-016-0685-0
11. Berisha V (2019) AI as a threat to democracy: towards an empirically grounded theory AI as a threat to democracy: towards an empirically grounded theory. Visar Berisha Autumn 2017 Supervised by Professor Joakim Palme, no. December 2017, pp 0–64
12. Kurke A, Smith L, Kim H (2020) The human right to democratic control of artificial intelligence contributors
13. Zekos G, (2022) How will AI influence. October 2022
14. Rosenbach E, Mansted K (2018) Can democracy survive in the information age?. Belfer Cent Sci Int Aff 1–22. [Online]. Available: https://www.belfercenter.org/publication/can-democracy-survive-information-age

15. Coeckelbergh M (2022) Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence. AI Ethics 0123456789. https://doi.org/10.1007/s43681-022-00239-4
16. Arana-Catania M et al. (2021) Citizen participation and machine learning for a better democracy. Digit Gov Res Pract 2(3). https://doi.org/10.1145/3452118
17. Zummo ML (2020) Performing authenticity on a digital political stage. Iperstoria 15:96–118
18. Mehr H (2017) Artificial intelligence for citizen services and Government. Harvard Ash Cent Technol Democr August, pp 1–16. [Online]. Available: https://ash.harvard.edu/files/ash/files/artificial_intelligence_for_citizen_services.pdf
19. Nemitz P (2018) Constitutional democracy and technology in the age of artificial intelligence. Philos Trans R Soc A Math Phys Eng Sci 376(2133). https://doi.org/10.1098/rsta.2018.0089
20. Wehsener A, Zakem V, Miller MN (2022) Future digital threats to democracy: trends and drivers
21. Danescu E (2021) Democracy, freedom, and truth at a time of digital disruption: an equation with three unknowns?. Fake News Is Bad News—Hoaxes, Half-truths Nat Today's J. https://doi.org/10.5772/intechopen.97662
22. Bernhard M, O'Neill D (2018) Digital politics. Perspect Polit 16(4):915–917. https://doi.org/10.1017/S1537592718003146
23. Feldstein S (2019) How artificial intelligence is reshaping repression. J Democr 30:40
24. Schneider I (2020) Democratic governance of digital platforms and artificial intelligence? Exploring governance models of China, US, the EU and Mexico. eJournal eDemocracy Open Gov 12(1):1–24. https://doi.org/10.29379/jedem.v12i1.604
25. Polyakova A, Meserole C (2019) Exporting digital authoritarianism: the Russian and Chinese models
26. York JC (2018) The impact of digital technology upon democracy. Japan SPOTLIGHT, December
27. Vīķe-Freiberga, V (2019) Digital transformation and the future of democracy. [Online]. Available: http://www.clubmadrid.org/digital-transformation-and-the-future-of-democracy-how-can-artificial-intelligence-drive-democratic-governance/
28. Iosifidis P, Nicoli N (2021) Digital democracy, social. Routledge, New York
29. Montgomery KC (2018) Youth and digital democracy: intersections of practice, policy, and the marketplace. Civ Life Online September. https://doi.org/10.7551/mitpress/7893.003.0003
30. Helbing D (2019) In: Towards digital enlightenment: essays on the dark and light sides of the digital revolution
31. Falque-Pierrotin I (2017) How can humans keep the upper hand? The ethical matters raised by algorithms and artificial intelligence. [Online]. Available: https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf

Analysis of Smoking Hazard Education Using Facebook Social Media: A Case Study of High School Students in Special Region of Yogyakarta, Indonesia

Kusbaryanto and Fairuz

Abstract This study aims to explain how smoking hazard education uses social media Facebook on high school students in the special region of Yogyakarta. This study used a Quasy experimental design with a non-equivalent control group design. The results of statistical tests showed that in the control group, the level of knowledge about the dangers of smoking obtained was p-value = 0.011 ($p > 0.05$), which was not significant. Meanwhile, the statistical test for attitudes about the dangers of smoking was p-value = 0.004 ($p < 0.05$), which was significant. The results of the statistical test in the experimental group showed that the level of knowledge about the dangers of smoking was p-value = 0.001 ($p < 0.05$), which was significant, while the statistical test for attitudes about the dangers of smoking was p-value = 0.001 ($p < 0.05$), which was also significant. The conclusion is education about the dangers of smoking in social media Facebook is effective in increasing the knowledge and attitudes of teenagers about the dangers of smoking among high school students in special region of Yogyakarta. It is expected that by conducting this research, teenagers can avoid the dangers of smoking. Thus, the implementation of education using Facebook social media needs to be disseminated in the community, especially among teenagers.

Keywords E-health · Facebook · Social media · Health education · Smoking hazard

Kusbaryanto (✉)
Master of Hospital Adminstration, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: kusbaryanto@umy.ac.id

Fairuz
Faculty of Medicine and Health Sciences, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_86

1 Introduction

Tobacco has been a fundamental part of human civilization since prehistory. About America, tobacco was first grown in 6000 BC. The hazards of smoking tobacco were made abundantly clear and widely acknowledged in the middle of the twentieth century, despite the official measures to stop smoking being approved or embraced. It is difficult to realize that when the Royal Australian College of General Practitioners (RACGP) was founded, medical experts, including general practitioners, were advising patients to stop smoking. An excellent success story for Australian public health was the drop in the age-standard daily smoking rate from 25.6% in 1989–1990 to 14.7% in 2016–17, which reflected the decline in smoking rates (2014–15). However, among Aboriginal and Torres Strait Islander people, smoking rates remain high. Although it is still very high, the smoking rate among Aboriginal and Torres Strait Islander persons has decreased from 50% in 2004–05 to 40% in 2018–19 [1].

At 2018 data reveals that 69% of US citizens currently use social media, with daily usage among Facebook users reaching 74% as people spend more and more time on it. Social media interventions have the ability to reach a huge number of smokers who are interested in quitting. Participants in this intervention were placed in exclusive social media groups (e.g., Facebook, Twitter). Several interventions publish content to their social media accounts (e.g., Facebook pages). Participants can interact with intervention content and each other simultaneously on social media platforms because they are designed to promote communication [2]. Enabling communication and support using modern platforms like Facebook, Instagram, and WhatsApp, where smokers who can stop do so can share their experiences. Providing messaging and support using modern platforms like Facebook, Instagram, and WhatsApp where smokers who have successfully stopped share details about their experience and how they are adjusting. In order to create a Facebook page for World No Tobacco Day 2016, tobacco care experts collaborated with a number of young medical students. They wrote posts on the topic of smoking and urged viewers to comment. The team employed Facebook advertising tools in the early stages of the campaign to promote the page, and these were funded by a number of supporters. A team is in charge of running the Facebook page and coming up with fresh concepts to broaden the audience and impact of the campaign. The website has long been popular on social media. For a number of months, the page has been trending on social media. 500 K people responded to the promotion, reaching 3 M people. Following the success tales on the page, about 3000 smokers were able to stop [3].

As one of the leading causes of disability and death worldwide, including in Turkey, smoking needs to be reduced. In the world, 27.1% of adults and 8.4% of young people smoke frequently, according to the 2012 Global Smoking Survey. It is emphasized that developing countries like Turkey are experiencing a dramatic increase in the prevalence of youth smoking. Studies show that the majority of smokers in various countries start their habit before the age of 18. One of the factors that leads young people to start smoking is the influence of their peers, parents, or siblings. Anti-smoking campaigns in the education sector have been successful in lowering tobacco use [4]. There is a behavioral aspect to smoking that is connected to a physical nicotine addiction. The numerous behavioral therapy methods that are accessible include individual behavior counseling, brief guidance/interventions, telephone counseling, open group forms of behavior therapy, and closed group forms of behavior therapy, to name a few. Open group behavioral treatment is superior and more affordable than other types, nevertheless. There is evidence that dental school student's attitudes and behaviors toward oral hygiene are suitable, and that they gain knowledge while attending dental school. The conclusions of other research suggesting health workers should get training on smoking prevention were supported by participants in our study [5]. With order to increase high school student's understanding of and attitudes toward the risks associated with smoking in Yogyakarta. This study looks into how well Facebook media social works for informing students about the dangers of smoking. The contribution of this research will assist adolescents in understanding the risks of smoking by providing this education.

2 Research Methods

This study uses quantitative methods. This study used a Quasy experimental design with a non-equivalent control group design. This sampling technique used purposive sampling with 32 respondents in the experimental group (24 men, 8 women) and 30 respondents in the control group (21 men, 9 women). Respondents were 10th grade students of SMU Muhammadiyah 7 Yogyakarta, and the research location was at SMU Muhammadiyah 7 Yogyakarta. The inclusion criteria were grade 10 students of SMU Muhammadiyah 7 Yogyakarta who followed the research to completion, while the exclusion criteria were students who had stayed in class. The data were analyzed using Wilcoxon and Mann Whitney, and the data were collected using a questionnaire.

3 Result

3.1 Control Group

As many as 60% of the respondents in the control group were 15 years old, and at least 6.7% were aged 17 years (Table 1).

In Table 2, the characteristics of respondents by sex show that the number of male respondents in the experimental group was 21 (70%), and the number of female respondents was 9 (30%) (Table 3).

In the control group, the data were tested using the Wilcoxon signed RanksTest because the data were not normally distributed. The test result on knowledge was $p = 0.475$, while the test result on attitudes was $p = 0.195$, which was not significant.

Table 1 Characteristics of control group respondents by age

Age (year)	Total	Percentage (%)
14	3	10
15	18	60
16	7	23.3
17	2	6.7
Total	30	100

Table 2 Characteristics of control group respondents by sex

Sex	Total	Percentage (%)
Male	21	70
Female	9	30
Total	30	100

Table 3 Statistical test results of knowledge and attitudes about smoking in the control group

Variable	Knowledge			Attitude		
	N	Mean	SD	N	Mean	SD
Pretest	30	15.53	2.11	30	63.67	12
Posttest	30	15.50	2.37	30	64.63	10.32
p		0.475**			0.195**	

* Significant ($p < 0.05$) ** Not significant ($p > 0.05$)

Thus, it can be concluded that there is no significant difference or no relationship because it is a significant value of 0.05.

3.2 *Experimental Group*

In Table 4, the characteristics of respondents based on age are divided into 4 groups. In this experimental group, 1 respondent (3.1%) was 14, 23 respondents were 15 (71.9%), 7 respondents were 16 (21.9%), and 1 respondent was 17 (3.1%).

In Table 5, the characteristics of respondents by sex show that the number of male respondents in the experimental group was 24 people (75%), while the female respondents were 8 (25%). In the control group, there were 21 (70%) men and 9 (30%) women (Table 6).

The results showed that in the experimental group, the knowledge value was $p = 0.001$ ($p < 0.05$), while the attitude value was $p = 0.001$ ($p < 0.05$). In the control group, the knowledge significance value was > 0.05, and the attitude significance value was > 0.05. These results indicated that in the experimental group, there was a

Table 4 Characteristics of the experimental group respondents by age

Age (year)	Total	Percentage (%)
14	1	3.1
15	23	71.9
16	7	21.9
17	1	3.1
Total	32	100

Table 5 Characteristics of the experimental group respondents by gender

Gender	Total	Percentage (%)
Male	24	75
Female	8	25
Total	32	100

Table 6 Results of statistical tests of knowledge and attitudes about smoking in the treatment group

Variable	Knowledge			Attitude		
	N	Mean	SD	N	Mean	SD
Pretest	32	13.53	2.44	32	62.38	12.89
Posttest	32	16.91	2.20	32	72.44	5.42
p		0.001			0.001	

* Significant ($p < 0.05$) ** Not significant ($p > 0.05$)

significant difference between before and after being given counseling. Meanwhile, in the control group, there was no significant difference. When comparing both groups, in the control group, the analysis results of knowledge and attitude obtained were $p > 0.05$, which was not significant, while in the treatment group, the results both in knowledge and attitude were $p < 0.05$, which was significant.

4 Discussion

Tobacco has been a fundamental part of human civilization since prehistory. About America, tobacco was first grown in 6000 BC. The hazards of smoking tobacco were made abundantly clear and widely acknowledged in the middle of the twentieth century, despite the official measures to stop smoking being approved or embraced. It is difficult to realize that when the Royal Australian College of General Practitioners (RACGP) was founded, medical experts, including general practitioners, were advising patients to stop smoking. An excellent success story for Australian public health was the drop in the age-standard daily smoking rate from 25.6% in 1989–1990 to 14.7% in 2016–17, which reflected the decline in smoking rates (2014–15). However, among Aboriginal and Torres Strait Islander people, smoking rates remain high. Although it is still very high, the smoking rate among Aboriginal and Torres Strait Islander persons has decreased from 50% in 2004–05 to 40% in 2018–19 [1].

At 2018 data reveals that 69% of US citizens currently use social media, with daily usage among Facebook users reaching 74% as people spend more and more time on it. Social media interventions have the ability to reach a huge number of smokers who are interested in quitting. Participants in this intervention were placed in exclusive social media groups (e.g., Facebook, Twitter). A number of interventions publish content to their social media accounts (e.g., Facebook pages). Participants can interact with intervention content and each other simultaneously on social media platforms because they are designed to promote communication [6]. Enabling communication and support using modern platforms like Facebook, Instagram, and WhatsApp, where smokers who can stop do so can share their experiences. Providing messaging and support using modern platforms like Facebook, Instagram, and WhatsApp where smokers who have successfully stopped share details about their experience and how they are adjusting. In order to create a Facebook page for World No Tobacco Day 2016, tobacco care experts collaborated with a number of young medical students. They wrote posts on the topic of smoking and urged viewers to comment. The team employed Facebook advertising tools in the early stages of the campaign to promote the page, and these were funded by a number of supporters. A team is in charge of running the Facebook page and coming up with fresh concepts to broaden the audience and impact of the campaign. The website has long been popular on social media. For a number of months, the page has been trending on social media. 500 K people responded to the promotion, reaching 3 M people. Following the success tales on the page, about 3000 smokers were able to stop [7].

In South Africa, adults smoked at a rate of 21.5% in 2016. About 20% of pulmonary tuberculosis deaths and 8% of all deaths in South Africa are attributed to smoking. Some of the top 10 deaths in South Africa, such as tuberculosis, pneumonia, heart disease, cerebrovascular disease, diabetes, hypertension, and chronic respiratory disorders, are caused by smoking or made worse by it. Quitting smoking lowers the risk of smoking-related illness and mortality. Although South Africa offers some professional resources (medication and counseling) and national stop lines to assist smokers, there are still access and utilization issues. Only 29.3% of smokers in South Africa received advice to stop smoking from medical professionals in 2012. South Africa has steadily rolled out policies to encourage smoking cessation over the past few decades [8].

As one of the leading causes of disability and death worldwide, including in Turkey, smoking needs to be reduced. In the world, 27.1% of adults and 8.4% of young people smoke frequently, according to the 2012 Global Smoking Survey. It is emphasized that developing countries like Turkey are experiencing a dramatic increase in the prevalence of youth smoking. Studies show that the majority of smokers in various countries start their habit before the age of 18. One of the factors that leads young people to start smoking is the influence of their peers, parents, or siblings. Anti-smoking campaigns in the education sector have been successful in lowering tobacco use [9]. There is a behavioral aspect to smoking that is connected to a physical nicotine addiction. The numerous behavioral therapy methods that are accessible include individual behavior counseling, brief guidance/interventions, telephone counseling, open group forms of behavior therapy, and closed group forms of behavior therapy, to name a few. Open group behavioral treatment is superior and more affordable than other types, nevertheless. There is evidence that dental school student's attitudes and behaviors toward oral hygiene are suitable, and that they gain knowledge while attending dental school. The conclusions of other research suggesting health workers should get training on smoking prevention were supported by participants in our study [10]. With order to increase high school student's understanding of and attitudes toward the risks associated with smoking in Yogyakarta. This study looks into how well Facebook media social works for informing students about the dangers of smoking. The contribution of this research will assist adolescents in understanding the risks of smoking by providing this education.

Respondents who did not smoke knew more about the doctor's awareness of smoking-related issues. More nonsmokers are accurately informed on how smoking can cause heart and lung disease in adults and children, both actively and passively. Furthermore, a greater proportion of nonsmokers are correct in their perception of the negative effects that passive smoking can have on children's health, particularly with regard to lower respiratory tract disease ($P = 0.016$) and infant death ($P = 0.013$). On average, 72% of nonsmokers and 54.5% of smokers ($P = 0.002$) agreed that fetal disease risk was raised by mother smoking during pregnancy [11]. Implementing these instructions in practice is still challenging because nicotine dependence is a chronic relapsing condition that requires continual effort to prevent relapse. Even though in many nations more than half of smokers aspire to stop and a third have made at least three attempts, less than half of smokers quit before the age of 60. Some of the barriers to intervention have been discussed, including a lack of knowledge, negative attitudes among medical professionals, low self-efficacy, inadequate training, competing priorities and the idea that counseling is an inappropriate service, time, energy, and resource constraints, a lack of skills, and worries about the doctor-patient relationship and the patient's insufficient motivation. Healthcare personnel who smoke are more likely than those who do not to dissuade patients from quitting in various countries. Healthcare professionals also claim that they lack trust in smoking cessation programs and a practical understanding of counseling techniques for quitting smoking. The biggest barrier to providing smoking cessation services between these standards is the inadequate training of healthcare professionals [12].

Statistics from 2018 show that from 5% in 2005 to 69% in 2018, more adults in the United States used at least one social media platform. Over the past two decades, internet usage has increased dramatically, and it now plays a significant role in shaping social culture. Social media has the potential to reach vast groups and offer low-cost interventions for positive behavior change. Social media makes it possible to offer social support and has emerged as a crucial tool for spreading awareness of the negative health impacts of smoking, teaching people how to quit, and changing attitudes about smoking-related behavior. As a result, numerous social media techniques have been actively accepted and utilized in recent years for smoking cessation treatments. Social media can be used to spread information about smoking to the general public, as well as to let smokers and counselors connect online [13]. The adoption of social media and mobile communications technologies as culturally relevant smoking cessation programs for the young adult Spanish-speaking population in South Texas is discussed in this paper. The ability to offer support services for quitting smoking via mobile devices is quite promising. According to a Cochrane study, texting or instant messaging greatly boosts the likelihood of successfully quitting smoking, with an average odds ratio of 1.7. Randomized experiments have demonstrated that telephone smoking cessation counseling protocols based on social cognitive theory, transtheoretical models, and motivational interviews greatly increase the effectiveness of smoking cessation, especially among young adults. Successful SMS distribution techniques can be modified for social media delivery in interactive chat applications, enabling help to be communicated through amusing and educational media material wherever the user may be. Latinos and other low-income young smokers who primarily communicate via mobile devices have a potential market for social media interactive chat technologies [14].

Distribution of smoking cessation behavior change stages among outpatients. In the stage of transformation, outpatients are split into groups based on whether they "have intentions to take action within 6 months" or "have no plans to take action in the next 6 months" (pre-contemplation). A total of 10.0% of the population will "start taking action within the next month" (preparation), 7.3% will "have taken action but not more than 6 months" (action), 8.7% will "have taken action for more than 6 months" (maintenance), and 5.3% will "begin taking action within the next month" (action) [15].

The vascular surgeons and stakeholders we spoke with believed that a swift but compassionate intervention with suggestions for counseling and straightforward medicine for willing patients formed the critical elements of a successful patient smoking cessation program. According to our patient interviews with smokers and ex-smokers, it is crucial to concentrate and carry out intensive interventions on brief, specific, and repeated therapies. Even while stakeholders and patients agree that the empathic approach is an essential part of smoking cessation programs, patients prefer the personalized nature and patient-specific scheduling of the strategy.

Smoking greatly raises the risk of mortality and morbidity in the US. The five A's framework is a tool that doctors can use to motivate their patients to quit smoking (ask, advise, assess, assist, arrange). Every time a patient visits the doctor, their tobacco usage should be brought up and their motivation to quit should be assessed. Doctors should strongly counsel their patients to quit smoking, and if they are still hesitant, they should use motivational interviewing techniques. It is crucial to underline the benefits and importance of quitting smoking, the risks of smoking, and anticipated barriers to abstinence in professional discussions with unmotivated patients. It is important to emphasize these points whenever possible. Individuals should receive appropriate pharmacological treatment in quitting, such as nicotine replacement therapy, bupropion, and varenicline. When using medication to assist with quitting, the success rate can be boosted by 50% [15]. Nearly 40 million adults smoke in the

US, and smoking is still a significant cause of avoidable illness and death. Smoking is connected to about 20 different types of cancer, including those that affect the liver, esophagus, stomach, lung, kidney, and bladder. Smoking is also to blame for a third of cancer deaths in the nation. Furthermore, smoking has been connected to a number of chronic disorders that impact almost every organ in the body, including diabetes, blindness, and cardiovascular and respiratory diseases.

Ten patients (6 men and 4 women) and 5 carers (all women) participated in the study. All of the patients smoked, with patients older than the average age of 41.7 years smoking a mean of 14 cigarettes per day. The average age of the patients was 59.4, and 2/5 of the nurses smoked. The following four main themes emerged: Most people are unaware of how smoking continues to affect cancer treatment; Many cancer patients do not always want to stop smoking; Previous failures to quit smoking can make patients feel hopeless about future attempts; Some people believe that smoking cessation treatment is not available at the time of cancer diagnosis or during cancer treatment17. The most significant adverse outcomes of antenatal cigarette smoking for both mother and child include preterm birth, placental abnormalities, low birth weight, perinatal death, and sudden infant death syndrome. The assessment of a woman's smoking status and the provision of advice and support for stopping smoking are crucial parts of prenatal care. The statistics suggest that pregnant women should stop smoking more frequently and smoke fewer cigarettes each day. When pregnant, 21% of Australian women stopped smoking, and 46% reduced their smoking. The actions and attitudes of medical staff members working in prenatal clinics regarding smoking evaluation and cessation advice have not yet been the subject of investigations in Pakistan.

5 Conclusion

The results of statistical tests showed that in the control group, the level of knowledge about the dangers of smoking obtained was p-value = 0.011 ($p > 0.05$), which was not significant. Meanwhile, the statistical test for attitudes about the dangers of smoking was p-value = 0.004 ($p < 0.05$), which was significant. The results of the statistical test in the experimental group showed that the level of knowledge about the dangers of smoking was p-value = 0.001 ($p < 0.05$), which was significant, while the statistical test for attitudes about the dangers of smoking was p-value = 0.001 ($p < 0.05$), which was also significant. The conclusion is education about the dangers of smoking in social media Facebook is effective in increasing the knowledge and attitudes of teenagers about the dangers of smoking among high school students in special region of Yogyakarta. It is expected that by conducting this research, teenagers can avoid the dangers of smoking. Thus, the implementation of education using Facebook social media needs to be disseminated in the community, especially among teenagers.

References

1. Dinh PC, Schrader LA, Svensson CJ, Margolis KL, Silver B, Luo J (2019) Smoking cessation, weight gain, and risk of stroke among postmenopausal women. Prev Med (Baltim) 118:184–190
2. Thrul J, Tormohlen KN, Meacham MC (2019) Social media for tobacco smoking cessation intervention: a review of the literature. Curr Addict Reports 6(2):126–138
3. Elmeguid WA, Kassem A, Abdalla R, Moustafa O (2018) Promoting smoking cessation through new media tools Facebook, Instagram and WhatsApp. J Glob Oncol
4. Agaku I, Egbe C, Yusuf OA (2021) Utilisation of smoking cessation aids among South African adult smokers: findings from a national survey of 18 208 South African adults. Fam Med Commun Heal 9(1)
5. İçmeli OS, Türker H, Gündoğuş B, Çiftci M, Aktürk UA (2016) Behaviours and opinions of adolescent students on smoking. Tuberk Toraks 64(3):217–222
6. Mostafa N, Momen M (2017) Effect of physician's smoking status on their knowledge, attitude, opinions and practices of smoking cessation in a University Hospital, in Egypt. J Egypt Public Heal Assoc 92(2):96–106
7. Hasan SI, Hairi FM, Tajuddin NAA, Nordin ASA (2019) Empowering healthcare providers through smoking cessation training in Malaysia: a preintervention and postintervention evaluation on the improvement of knowledge, attitude and self-efficacy. BMJ Open 9(9):e030670
8. Luo T, Li MS, Tseng TS (2021) Using social media for smoking cessation interventions: a systematic review. Perspect Public Health 141(1):50–63
9. Chalela P, McAlister, AL, Amelie G, Ramirez AG (2022) Facebook chat application to prompt and assist smoking cessation among Spanish-speaking young adults in south texas. Health Promot Pract 23(3):378–381
10. Hsu CY, Liao HE, Huang LC (2020) Exploring smoking cessation behaviors of outpatients in outpatient clinics: application of the transtheoretical model. Med (Baltimore) 99(27)
11. Newhall K, Burnette M, Brooke BS, Schanzer A, Tan T, Flocke S, Farber A, Goodney P (2016) Smoking cessation counseling in vascular surgical practice using the results of interviews and focus groups in the Vascular Surgeon offer and report smoking cessation pilot trial. J Vasc Surg 63(4):1011–1017
12. Larzelere MM, Williams DE (2012) Promoting smoking cessation. Am Fam Phys 85(6):591–598
13. Potter LN, Lam CY, Cinciripini PM, Wetter DW (2021) Intersectionality and smoking cessation: exploring various approaches for understanding health inequities. Nicotine Tob Res 23(1):115–123
14. Barrett JR, Stafford LC, Alagoz E, Piper ME, Cook J, Flohr SC, Weber SM, Winslow ER, Kelly SMR, MD, Abbott DE (2019) Smoking and gastrointestinal cancer patients—is smoking cessation an attainable goal?. J Surg Oncol 120(8):1335–1340
15. Ghazal S, Akhter S, Ali U, Rizvi N (2017) Knowledge of female doctors about smoking risks and their attitude toward cessation in antenatal clinics-perspective from tertiary care hospitals in Karachi. J Pak Med Assoc 67:1809–1813

Analysis of Infotainment Programs in Digital Media: Legal Protection for Indonesian Children Perspective

Nanik Prasetyoningsih and Moli Aya Mina Rahma

Abstract Infotainment programs in Indonesian digital media are currently criticized by social media users, as the content of the shows is often not appropriate for child viewers. Whereas children can now access and watch shows easily using social media. Therefore, this research aims to explain the regulations that apply in the dissemination of infotainment programs on digital media and what further efforts are possible to achieve legal protection for children from infotainment programs on digital media. This qualitative research found that the Broadcasting Law and Broadcasting Code of Conduct and Broadcast Program Standards have regulated the protection of children from infotainment programs in digital media. Child protection is also specifically regulated in the Child Protection Law. Then regarding the content of information disseminated in digital media is regulated in the Electronic Information and Transaction Law. It is critical that the government enacts new legislation to protect children who watch infotainment shows, both those that are aired and those that are accessible on social media.

Keywords Media digital · Infotainment program · Social media · Legal protection · Children

1 Introduction

The infotainment program in Indonesia is currently receiving criticism from social media users because it usually highlights celebrity personal struggles that have absolutely nothing to do with the general public. This infotainment show is regarded as one of the factors contributing to people's popularity, not because of their accomplishments, but rather because of their sensations. Children may now access and watch infotainment programs; thus, this is an important factor to think about. In Indonesia, people consider television to be one of the most dependable electrical

N. Prasetyoningsih (✉) · M. A. M. Rahma
Master of Law, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: nanikprasetyoningsih@umy.ac.id

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_87

devices. Technology use has advanced at an incredible rate during the past few years [1].

Children are using smartphones more frequently these days, to the point that it could lead to psychological issues among children. Therefore, the risks associated with smartphone use must be understood. Minors are especially vulnerable since they struggle with self-control and have underdeveloped control skills [2]. Children typically imitate what they see on television or social media, including infotainment, which is detrimental to their development and is frequently presented on digital media. However, what is presently portrayed in entertainment is the private lives of prominent leaders, which should not be exposed. Due to the tremendous degree of client interest in this event, this is done regularly. reveals that infotainment shows are no longer suitable for viewing, especially by youngsters under the age of 18, and that greater efforts are needed to address this issue.

Although using technology and social media can lead to many great outcomes, it can also be a harmful impact, particularly if children use them. There are more parents today are trying to prevent their kids from playing video games because they are aware of the harm that these games can do to kids. Additionally, modern parents may find it difficult to understand digital games because to its hybrid and sophisticated nature, which raises concerns about kids abusing digital media [3]. It can also have a comparable effect on a child's body and mind, encouraging them to continually think positively, actively, and imaginatively. Since families can use television to develop intimacy by watching together, its accessibility and the programs on it are predicted to have a good impact. However, the positive impact that the community had hoped for has not fully materialized. There currently needs to be a show on Indonesian television that differentiates between what is appropriate for adults and what is appropriate for children, and no adjusted broadcast hours are in place [4].

The free-market competition growth, in which much money finds easy to acquire local or small media, is one of the effects of globalization in this medium [5]. One of the reasons the owners will carry on with operations is the media's propensity to focus more on business. As a result of their enormous popularity with the general public and their importance as the media industry's main source of revenue, infotainment shows are still televised. Infotainment is the show that has attracted interest, as evidenced by the development of several programs' infotainment. In actuality, Indonesian entertainment programs merely present the same content every day, and even trivial things are made to seem like important facts that everyone should be aware of. Surprisingly, there are still a surprising number of people who can name viewers of the show who are oblivious of its significance or urgency.

2 Research Method

This research concerns to the law. Research is being done using a normative legal approach. Conducting legal research is an effort to find legal regulations, principles, and doctrines. Legal research is done to provide arguments, theories, or original perspectives that can be used as solutions to the problems at hand. Qualitative analysis methods are used to analyze textual and narrative data. The entire set of data is then programmed to meet user requirements. The method of data analysis is also applied. Actually, data interpretation and coding take place simultaneously. The classification and interpretation of data are carried out simultaneously. The process of interpretation serves the data to obtain the information needed.

3 Result and Discussion

3.1 Form the Law Governing Indonesian Broadcasting and Infotainment Programs

The reality of entertainment today is completely at odds with the laws that are already in place. Currently, news concerning celebrities' private life that has absolutely no elements of public interest is regularly presented on infotainment programs. Examining the news or information that is broadcast reveals that most celebrities today hardly have any privacy because all of their personal struggles are constantly made public to a big audience. However, in situations like these, it can be difficult to pinpoint exactly who is to blame because news in today's media is frequently dominated by celebrity-related personal issues as well as the celebrities themselves.

The establishment of the Indonesian Broadcasting Commission is one of the Government's initiatives to encourage high-quality broadcasts, which essentially entails regulating every program broadcast. The Indonesian Broadcasting Commission is a body having considerable control over managing television shows or broadcasts at both the national and local levels [6]. The Indonesian Broadcasting Commission is in charge of overseeing the enforcement of regulations, Broadcasting Behavior Guidelines, and Broadcast Program Standards. Moreover, the Indonesia Broadcasting Commission has the authority to impose sanctions for the violation of the Broadcasting Behavior Guidelines and Broadcast Program Standards on television programs [7]. The Indonesia Broadcasting Commission has several authorities, such as: setting broadcast program standards, drafting regulations and establishing a broadcasting code of conduct, monitoring the application of regulations and guidelines of broadcasting, as well as the broadcast program standards, giving penalties for violations of broadcasting regulations and guidelines and broadcast program standards.

Recently, the performance of Indonesia Broadcasting Commission was questioned because some TV stations and programs that received warnings, but some other TV station who has the same programs from, there did not get reprimands [8]. Therefore, it needs to be criticized where the Indonesia Broadcasting Commission firmness toward similar shows with the same content. It seems that this institution is not firm in enforcing the rules regarding broadcasting to TV stations that have committed violations. Indonesia Broadcasting Commission must be able to dissuade rule offenders by having the courage to impose punishments other than a warning [8].

Nowadays, the infotainment shows mentioned above are not only broadcast on television but also through other platforms such as YouTube, Instagram, TikTok, and Facebook. The problem is that these platforms' infotainment content was outside the control of the Indonesian Broadcasting Commission. However, until now, there has been no independent institution that supervises social media content such as Indonesian Broadcasting Commission in supervising television and radio. Social media has evolved into an online conversation in which people create, share, bookmark, and network at an astounding rate [9]. As an internet-based (online) social interaction medium, social media allows its users to share, engage, and produce varied content in the form of blogs, wikis, forums, and social networks. [10]. The use of social media for news consumption is a two-edged sword [11]. On the one hand, individuals seek out and consume news via social media due to its low cost, ease of access, and rapid transmission of information [12]. On the other side, it facilitates the widespread dissemination of "fake news" [13], low-quality news containing purposely misleading material [11]. Furthermore, the negative impact of social media increases the distance between close people and vice versa; face-to-face interactions tend to decline, as well as make people addicted to the internet, generating problems and privacy issues, which are readily influenced by the ill influence of others [12]. Social media can also have a bad influence on children; moreover, the ease with which children use social media can cause them to spend longer time using social media [14].

The government has already passed legislation restricting the use of social media and electronic media, specifically the Law of Information Technology and Electronic in 2016. The role of the government regarding the children's protection from infotainment from these platforms was to supervise the electronic media for using electronic information and documents and providing facilitate information technology and electronic transactions uses. Furthermore, the government is required to protect the public interest from all types of effects caused by the misuse of electronic information and transactions that may disrupt public order in accordance with the provisions of laws, and for this purpose, the government can prohibit the dissemination and use of electronic information and documents containing prohibited contents through the provisions of laws.

To carry out the prevention, the government could be entitled to terminate access or direct the electronic system operator to discontinue access to electronic information and documents containing illegal material.

As a result, in this scenario, the government has the ability and duty to regulate the manner of electronic information transmission in order to maintain public order. However, due to their early age, children and adolescents may meet problems during their searches, requiring parental interaction ensures they use reliable web resources, adequately absorb the material, and stay calm by the information they are reading. Inquiring about their children's and teen's online searches can aid in discovering and discussing this material [15]. Most adults are unaware that adolescents are particularly vulnerable to the threats posed by social media. The greatest dangers are peer-to-peer; incorrect information; a need for understanding online privacy issues; and third-party advertising group impacts [15].

3.2 The Indonesian Law Protecting Children from Infotainment Programs

Today's children and teens are surrounded by a digital environment such as television (TV), radio, and periodicals, which have been furnished with modern digital technologies that stimulate interactive and social connections and allow instant access to children and teens for fun, information, and knowledge; interpersonal interactions [16]. On the bright side, technology allows youngsters to play, explore, and learn in some ways. This learning opportunity is a critical developmental phase in children and involves the study of nature and the discovery of their own environment since children's brains are particularly malleable throughout this period [17], and the negative side, children's media can create cognitive impairment and diminish cognitive capacity [18].

Child protection is expressly controlled under Law of Children Protection. Children's protection is necessary and must be ensured to improve children's quality of life as they grow and develop. Parents, families, the community, the state, the government, and local government must be guaranteed, protected, and fulfilled the Children's rights [19]. According to the Law of Children Protection, child protection aims to: protect and guarantee children's rights in order for the children can live, grow, and develop; protect and guarantee children's rights to participate through human dignity; and protect children from inhumanity and intolerance while granting them special consideration while granting them special consideration. Moreover, children protection principles according to the 1945 Constitution are follows: no-discrimination policy, children's best interests, the right to live, to survive, and to develop, and value the child's viewpoint.

Obtaining infotainment is one of the children's information rights, but it must be supported with parental engagement to assist them in accessing positive news and information that will benefit them in the future because they are in the same environment as their children [20], as well as society. Parents play the most crucial role in guaranteeing this. The government is responsible for protecting children as consumers of television broadcasts such as infotainment by ensuring that what children see or watch is valuable information for them [21]. Children are given the best chance to grow and develop by being cared for both physically and mentally by creating a sense of security and comfort, creating an appropriate atmosphere, protecting children from consuming harmful substances, and not allowing them to consume harmful substances for their developing.

As stated by Article 17 of the United Nations Convention on the Rights of the Child, participating countries recognize the importance of the media's role and will make certain that children have access to information and resources from various national and international sources, particularly those aimed at promoting their social, mental, and moral well-being and physical and mental health. Therefore, UN member states will: (1) encourage the dissemination of mass media information and materials that are useful with the spirit of the children socially and culturally; (2) encourage international cooperation in the formation, information interchange and transmission from many cultures, both national and international; (3) encourage the creation and distribution of children's books; (4) encourage the media to pay special attention to the language requirements of minority and indigenous children; (5) encourage the establishment of adequate rules to ensure children's information; and materials that are detrimental to the welfare of the child, considering the requirements of Articles 13 and 18; and (6) protection of a child has become a policy for countries that are members of the United Nations (UN).

The Children Protection Law's was ensuring the protection of children from broadcasting programs, as stated that children are special audiences consisting of children and adolescents who are not yet 18 (eighteen) years old. The children protection is one of the foundations for developing Broadcast Program Standards from the Broadcasting Code of Conduct. Broadcast Program Standards 2012 states that the institutions of broadcasting are required to offer protection and empowerment to children by transmitting broadcast programs at the appropriate time in line with broadcast program categorization, and shall consider children's interests in all aspects of broadcast production, such that youngsters cannot access adult programs that will be transmitted after 10 PM. local time.

Under Broadcast Program Standards, children protection precautions are also mandated requirement for broadcast programs, as follows: (1) children's and/or teenagers' interests must be considered and protected in broadcast programming, (2) broadcast programs containing immoral content and/or information about allegations of immoral crimes are prohibited from displaying children and/or teenagers, (3) broadcast programs showing children and/or teenagers in events/law enforcement must conceal their faces and identities, and (4) live broadcast shows featuring minors are not permitted to air after 9.30 PM. local time.

The Indonesian Broadcasting Commission, in the fourth point which was for the protection of children and adolescents by paying attention to the availability of programs for children from 5 AM. until 6 PM. with content, storytelling style, and appearance in accordance with the psychological development of children and adolescents. Selectively selecting broadcast content to not encourage youngsters to mimic or consider common/ordinary activities that have lately been publicized, such as marriage at a young age, exploitation of early marriage, revelation of home issues, and so on. Confrontations, violent acts/scenes, and bullying in households, schools, and other social contexts and Keep scenes of love and infidelity to a minimum. The circular letter is one type of action done by the government, so all broadcasting institutions are encouraged to follow it.

As a result, every show broadcast must have acceptable substance, narrative style, and look that does not aggravate children's development and health [22]. There are still numerous situations and tales, as well as the presentation of infotainment shows, that are inappropriate for children and harmful to their development and health. It is fairly unusual for the information delivered by the host in the infotainment shows to be provocative. In light of all broadcasting rules, it is possible to infer that the show must not contain anything that encourages youngsters to learn about harmful behavior. Things like courting are improper. Currently, celebrity romance stories dominate infotainment broadcasts as if they were a typical occurrence. Furthermore, it is forbidden for infotainment programs to offer information that encourages children to learn about improper behavior and/or to rationalize inappropriate behavior as a normal part of everyday life. This implies that great care must be taken to protect children, teens, and women [23].

3.3 Legal Efforts Can Be Made in the Future to Protect Children from Infotainment Programs Better

The government needs to put more effort into creating engaging and high-quality informational content for children in order to strengthen legal protection. Legal protection aims to provide young viewers of material on television or in other media with a sense of legal security. The state provides legal protection in the form of acknowledged legislative standards that participants in the broadcasting industry are required to meet and implement. To ensure that every kid has the opportunity to exercise their natural, physical, mental, and social development and maturation.

Infotainment is rarely used for news shows that contain actual information, and they are better known as entertainment news. The balance of news content must be supported by various accurate data, opinions, comments to public opinion through interviews. This definition is in stark contrast to the concept of an infotainment program which does not require actual and factual dimensions because the infotainment format actually seems one-way even though it is presented in an entertainment style packaging format, such as a scriptwriting standpoint, but the content of this program format only provides information or an explanation about a product.

For this reason, the media must be able to select the sort of quality news that will avoid harmful activities that would negatively affect young customers. Because of this, by designating another country as a reference, many steps may be taken to set stronger children's legal protection from infotainment programs. *First*, this should be a concern in terms of improving the quality of child development. One of the things that may be done is to create a new provision for infotainment shows during the emergency time, which is not only in the form of circulars but also regulated in Law of Broadcasting. One of the provisions is to shorten the length of newscasts. If an increase in the quality of infotainment shows cannot be accomplished, the shows' cancelation might be considered. *Second*, Categorize Infotainment Program. The first step in clarifying the spectacle/program is to categorize it. When organizing television broadcast programs, each broadcast program that will be broadcast must first be categorized since the classification determines how and when a program may be broadcast. So, in order to improve child protection, infotainment programs should be classified as shows for adults so that children cannot access infotainment shows on television, since it has been mentioned that there is virtually no content acceptable for children in infotainment shows at the moment.

As stated on Article 36 of Broadcasting Law, that broadcast content is required to provide protection and empowerment to special audiences, namely children and adolescents, by broadcasting events at the appropriate time, and broadcasting institutions are required to include and/or mention the classification of audiences based on the content of the broadcast. Followed by Broadcast Program Standards in Article 14 and 15 state that broadcasting institutions must pay attention to the interests of children in all aspects of broadcast production, and protect the interests of children and/or youth.

This is a form of child safety measure. Children will understand that there are certain shows that are appropriate for their age and others that are not thanks to the age restrictions on each program. Furthermore, infotainment broadcasts after 10 PM are one of the important efforts because children do not comprehend how television programs are categorized based on rating categories. The previous article makes it apparent that the article's requirements for preventative measures must be met. Every program that will be aired must be prescreened to ensure that it is appropriate for children.

Social media must be closely monitored by a specialized organization, or else the government should impose restrictions. The distribution of content to children is subject to a number of regulations. This was previously requested by the Minister of Communication and Information in 2019, that this could be proposed, but it must be based on clear regulations, as well as the consideration that Indonesian Broadcasting Commission that must oversee the digital realm must also be reviewed first, because the Indonesian Broadcasting Commission's authority is currently limited to conventional media, namely television and radio, under the Law on Broadcasting.

4 Conclusion

It is critical that the government enacts new legislation to protect children who watch infotainment shows, both those that are aired and those that are accessible on social media. The government has passed numerous laws to safeguard children from entertainment, including the Law of Children Protection, the Law of Information and Electronic Transactions, and the Law of Broadcasting. The Broadcasting Code of Conduct, Broadcast Program Standards, and the Broadcasting Law supplied by the government must be followed by broadcasters. The government also founded the Indonesian Broadcasting Commission as an independent organization with authority to supervise media platforms like television and radio.

Since there is essentially no content suitable for children in infotainment shows at the moment, infotainment programs should be categorized as shows for adults in the future to promote child protection by preventing children from watching infotainment shows on television. In the meantime, the aforementioned infotainment programs are not only broadcast on television but also on other platforms like YouTube, Instagram, TikTok, and Facebook. Due to their youth, children and adolescents may encounter inaccuracies during searches; these platforms require parental engagement to ensure they are accessing credible online resources to evaluate the information and remain calm.

References

1. Choi K-S, Lee S-S, Lee JR (2017) Mobile phone technology and online sexual harassment among juveniles in South Korea: effects of self-control and social learning. Int J Cyber Criminol 11(1):110
2. Choi M, Tessler H, Kao G (2020) Arts and crafts as an educational strategy and coping mechanism for Republic of Korea and United States parents during the COVID-19 pandemic. Int Rev Educ 66(5):715–735
3. Dong PI (2018) Exploring Korean parents' meanings of digital play for young children. Glob Stud Child 8(3):238–251
4. Doly D (2017) Politik Hukum Pelindungan Anak Terhadap Program Siaran Televisi. Kajian 21(4):297–319
5. Yuniarto PR (2016) Masalah globalisasi di Indonesia: Antara Kepentingan, Kebijakan, dan Tantangan. J Kaji Wil 5(1):67–95
6. Ridwan M (2021) Peran KPI Dalam Proses Pengawasan Siaran TV Nasional di Indonesia. J Ilm Publipreneur 9(2):21–28
7. Kandyoh BV (2018) Sanksi Hukum atas Pelanggaran dalam Pembuatan Siaran Iklan Niaga Menurut Undang-Undang Nomor 32 Tahun 2002 Tentang Penyiaran. Lex Soc 6(6):43–63
8. Afifah A, Milla MN (2018) Penguatan Wewenang Komisi Penyiaran Indonesia sebagai Upaya Menurunkan Perilaku Pelanggaran Standar Penyiaran Televisi. Deviance J Kriminologi 2(1):1–17
9. Asur S, Huberman BA (2010) Predicting the future with social media. In: Proceedings—2010 IEEE/WIC/ACM ınternational conference web ıntelligence WI 2010, vol 1. pp 492–499
10. Istiani N, Islamy A (2020) Fikih Media Sosial Di Indonesia. Asy Syar Iyyah J Ilmu Syari'Ah Dan Perbank Islam 5(3):202–225
11. Granskogen T, Gulla JA (2017) Fake news detection: network data from social media used to predict fakes. CEUR Workshop Proc 2041(1):59–66
12. Cahyono AS (2016) Pengaruh Media Sosial Terhadap Perubahan Sosial Masyarakat Di Indonesia. Publicana 9(1):140–157
13. Juditha C (2018) Hoax communication ınteractivity in social media and anticipation (Interaksi Komunikasi Hoax di Media Sosial serta Antisipasinya). Pekommas 3(1)
14. Fitri S (2017) Dampak Positif Dan Negatif Sosial Media Terhadap Perubahan Sosial Anak. Nat J Kaji Penelit Pendidik dan Pembelajaran 1(2):118–123
15. O'Keeffe GS et al (2011) Clinical report—the impact of social media on children, adolescents, and families. Pediatrics 127(4):800–804
16. Chassiako R, Linda Y (2016) Children and adolescents, and digital media. Pediatrics 138(5)
17. Mustafaoğlu R, Zirek E (2018) The negative effects of digital technology usage on children's development and health. Addicta Turkish J 5(2)
18. Anderson DR, Subrahmanyam K (2017) Digital screen media and cognitive development. Pediatrics 140(Supplement 2):S57–S61
19. Browne KD, Hamilton-Giachritsis C (2005) The influence of violent media on children and adolescents: a public-health approach. Lancet 365(9460):702–710
20. Sanders MR, Montgomery DT, Brechman-Toussaint ML (2000) The mass media and the prevention of child behavior problems: the evaluation of a television series to promote

positive outcomes for parents and their children. J Child Psychol Psychiatry Allied Discip 41(7):939–948
21. Gentile DA, Saleem M, Anderson CA (2007) Public policy and the effects of media violence on children. Soc Issues Policy Rev 1(1):15–61
22. Brown JD, Childers KW, Bauman KE, Koch GG (1990) The influence of new media and family structure on young adolescent's television and radio use. Commun Res 17(1):65–82
23. Marinescu V (2008) Media covarage of 'Grasstoots' violence against women, a comparative analysis for Romania and Canada. Brazilian J Res 4(1):140–158

Personal Data Protection in Indonesian E-commerce Platforms: The *Maqasid Sharia* Perspective

Mızan Islami Nurzihad, Muchammad Ichsan, and Fadia Fitriyanti

Abstract Personal data must be kept secret because they are amanah (a trust) confided by the data owner to an authorized party. In e-commerce, large amounts of data must be kept secret for the safety of the data owner. Therefore, this research aimed to analyze the regulation of personal data protection in Indonesia and its compatibility with Maqasid Sharia (goals of Sharia). A normative approach and legal analysis method were employed. This research described the concept of personal data protection and observed that its regulation is contained in over 30 several laws and regulations in Indonesia. Based on their content, the laws could be divided into several categories, namely health, finance and business, human rights, state governance, and crime prevention. The result indicated that the draft and elements of the personal data protection law in Indonesia are compatible with the Maqasid Sharia for some arguments. This research recommended the establishment of an independent commission to protect the secrecy of data.

Keywords E-commerce · Indonesian regulation · *Maqasid Sharia* · Personal data protection

1 Introduction

The volume of trade transactions at Indonesian e-commerce companies previously increased by 5–10 times [1]. Although a fifty percent increase in new customers was recorded, the delivery or distribution of goods experienced delays due to transportation restrictions during the lockdown [2]. E-commerce refers to any form of trade transactions involving goods and services conducted through electronic media [3]. Trade via e-commerce may involve business-to-consumer (B2C) and business-to-business (B2B) transactions, as well as trade with structured electronic data exchange

M. I. Nurzihad · M. Ichsan (✉) · F. Fitriyanti
Master of Law, Universitas Muhammadiyah Yogyakarta, Postgraduate Program, Yogyakarta, Indonesia
e-mail: drichsan65@umy.ac.id

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_88

[3]. Despite the convenience of online transactions, including shopping, consumers must be cautious of data theft. E-commerce platforms require users to create an account with their personal information. Unfortunately, customer's data, such as names, e-mail addresses, phone numbers, and home addresses, may be breached, while the remaining data in the form of payment transaction information may be kept safe. This was experienced by Tokopedia users, particularly through financing digital OVO and credit cards [4]. A total of 13 million Buka Lapak account data were leaked on May 6, 2020 [5], while information of about 91 million users from the e-commerce platform Tokopedia was breached and spread on Internet forums [6]. In addition, 1.1 million Lazada accounts were leaked on November 2, 2020 [7].

In early 2020, research about personal data protection in Indonesia discussed the effort to protect citizen's information and the personal data protection bill. Due to their vast and obscure aspects, experts, scholars, privacy advocates, and research institutions opposed the adequacy of data protection laws [8]. Rosadi stated that a combination of regulations or a hybrid concept is the most appropriate regulatory concept. This hybrid is a concept that combines several approaches to regulating personal data privacy, particularly in e-commerce, and was chosen due to the rapid development of information technology, which allows information to be easily accessed, processed, compiled, and distributed [9]. Moreover, Sinaga emphasized that the legislation on personal data protection is still insufficient to protect consumers whose private information is leaked on the Internet [10]. Furthermore, Angriani stated that customers have rights and obligations to protect their data in Islamic and positive laws [11].

Consequently, this research complemented previous investigations on personal data protection by exploring its relationship with *Maqasid Sharia,* which was not covered by previous investigations. The problem of this research was to describe the extent personal data protection mechanisms on Indonesian e-commerce platforms follow *Maqasid Sharia.* It attempted to determine the compatibility between personal data protection in an e-commerce platform and *Maqasid Sharia* and is hoped to serve as a reference for future research.

This research was based on the argument that the personal data protection laws in Indonesia, both existing and in the form of bills, follow *Maqasid Sharia.* Most of the citizens in the country are Muslim, including governmental bodies from the executive, legislature, and judicial arms. As a result, Indonesian law naturally protects the interests of all citizens, including Muslims. *Maqasid Sharia* seeks to connect God's will with human goals or desires and allows the *ummah* (people) to play an essential role in developing and interpreting *maslahah* for humanity without departing from the essence of Islamic teachings.

2 Indonesian Personal Data Protection and Its Compatibility with Maqasid Sharia

Julie Innes (1992) defined privacy as a condition where a person controls their private decisions, including access, information, and actions [12]. Private is explained as a product of love, liking, and concern for others. This corresponds with the explanation by Solove (2008) that the context of privacy includes family, body, gender, home, communication, and personal information [12]. Gavison (1980) described privacy as a complex concept consisting of three independent and reducible elements, namely confidentiality, anonymity, and solitude. Each of these elements is independent, as their intrusion may cause a loss or violation [12].

The Ministry of Communication and Informatics of Indonesia defined personal data protection as protection during the acquisition, collection, processing, analysis, storage, appearance, announcement, delivery, dissemination, and/or opening of access, as well as the destruction of personal data [13]. In addition to the scope of personal data protection, which covers all aspects and stages of personal processing data, the *Permenkominfo* (Regulation of the Minister of Communication and Informatics) also regulates the rights of the data owner, as well as the obligations of data users and the electronic system operator at all stages of processing.

Since Indonesia has a Muslim majority, lawmakers create regulations for the welfare of all citizens, including Muslims. These regulations should follow *Maqasid Sharia* to achieve welfare. Nuruddin affirmed that *Maqasid Sharia* is made up of two words. They are *Maqasid,* the plural version of *maqsud,* which means intentional or purposeful, and *Sharia,* a linguistic term that refers to a path leading to water sources, supposedly considered the primary source of life [14]. Al-Syatibi stated that Sharia aims to realize benefits in the living world and hereafter and attested that the laws are prescribed for the benefit of the servant of God [15].

Allah created Islamic law (Sharia) with a specific purpose or goal in mind, which is to bring benefit or goodness to humans and protect them from harm or danger in this world and the next [16]. Based on these goals, Islamic law can be said to differ significantly from human-made law because the *maslahat* (benefit) is to be enjoyed and the *madharrat* (harm) prevented in the present and the future. Hence, followers of this law will receive good and avoid danger in both worlds [16].

Maqasid Sharia benefits are generally divided into three parts, namely *dharuriyat* (necessities), *hajiyat* (needs), and *tahsiniyat* (improvements). According to Al-Ghazali, *dharuriyat* is a collection of benefits that ensure the preservation of the five necessities, comprising *Hifz Din* (the right to religion), *Hifz Nafs* (the right to life), *Hifz 'Aql* (the right to education, think, have an opinion and press freedom), *Hifz Nasl* (reproductive, family, mother's, children's, civil, organizational, assembly, social, inheritance, and will rights or privileges), and *Hifz Mal* (economic, property, work, and worker's rights) [17].

Consequently, this research divided the existing laws and the bill regarding personal data protection into several categories, namely health, finance and business, human rights, state governance, and crime prevention. This classification was done based on their contents to determine the compatibility between the personal data protection laws in Indonesia with the *Maqasid Sharia.*

2.1 Health

Four laws regarding personal data protection can be classified into health categories. These are Law Number 29 of 2004 concerning Medical Practice, of which articles 51 and 52 state that medical practice is regulated to protect patients. Doctors and dentists are obliged to preserve the knowledge of treatment and ensure it is kept secret even after the patient's demise [18]. Law Number 36 of 2009 concerning Health specified that everyone has the right to obtain information about their health, including actions and treatment that have or will be received from health workers [19]. Law Number 40 of 2009 concerning Hospitals [20] and Law Number 36 of 2014 concerning Health Workers affirmed that every patient has the right to privacy and confidentiality [21]. Finally, Law Number 18 of 2014 concerning Mental Health demands that the *'aql* or mind must be protected [22].

These laws follow *Maqasid Sharia* in maintaining life, mind, and lineage. The Qur'an emphasized the preservation of human life as the word of Allah in surah Al-Maidah: 32, which reads: "*And whoever preserves the life of a human being, it is as if he preserves the life of all humans."* Respecting the right to life is a fundamental law, regardless of the person, position, or profession. The obligation of a person to protect the rights of other human beings is a sacred mission outlined in religion and international human rights treaties. It entails safeguarding personal as well as the general interests of many aspects of human life.

Islam also values reason very much. With reason, man thinks, develops, and discovers things that benefit his life in the world and the hereafter. The Islamic way of maintaining reason includes commanding humans to open and gain knowledge. Moreover, Islam forbids a man from corrupting his intellect by consuming the destroyer of reason, such as liquor, drugs, and others [16]. In addition, Islamic teachings command that offspring be kept so that there is human continuity on the face of this earth. The trick includes maintaining self-honor and getting married. It is because man obtains good offspring for these two reasons. For that reason, the maintenance of personal data contained in the above laws is in line with *Maqasid Sharia.* If personal data is not protected, it will have the potential to be misused, causing the honor of themselves and their descendants to be not maintained.

2.2 Finance and Business

Based on research, seven of the laws governing personal data protection are finance and business related. They include Law Number 10 of 1998 concerning Amendments to Law Number 7 of 1992 concerning Banking, Law Number 23 of 1999 concerning

Bank Indonesia, Law Number 21 of 2008 concerning *Sharia* Banking, and Law No. 21 of 2011 concerning Financial Service Authority, which have similar concepts that banks, in their capacity as depositors, are required to keep customer information confidential [23]. Based on this statement, the personal data protection laws in finance categories aim to protect their customers from any financial harm. This is compatible with *Maqasid Sharia* in protecting the wealth and property of bank customers.

In addition, Law Number 8 of 1997 concerning Company Documents states that the confidentiality of company documents must be maintained by the employer and employee [24]. Such documents may include elements of company management, profit and loss statements, and employee's personal data. Law Number 36 of 1999 concerning Telecommunications allows the telephone company to record customer conversation but only authorizes its release following the permission of the Attorney General or the head of the police force [25]. This regulation enables the government to monitor the actions of citizens in the country. Law Number 19 of 2016 concerning Amendments to Law Number 11 of 2008 concerning Information and Electronic Transactions explains that the protection of private data in using information technology is a part of the personal rights of an individual [26].

The above laws follow *Maqasid Sharia* in maintaining a citizen's wealth, dignity, and lineage. Islam views privacy as deserving of respect because it is related to one's confidentiality. Generally, banking activities in Islam are founded on Sharia principles derived from the Al-Quran. As stated in the letter Al Baqarah (2) verse 275, there is a prohibition on usury and the permissibility of buying and selling. The Qur'an and Sunnah describe four goals of Islamic banking activities based on Sharia. They are (1) Prioritizing Allah's worship above all else. (2) Creating Sharia bank activities to achieve a healthy life in the hereafter by obtaining heaven. (3) Providing a mechanism for distributing the funds of the rich to the needy. (4) Achieving the predetermined economic objectives. This means Islamic banking activities can impact all communities positively [27].

2.3 Human Rights

Only two laws regarding personal data protection were associated with human rights. They are Article 28 H paragraph (4) of the 1945 Constitution, the first law in Indonesia, which states that everyone is entitled to personal property rights, meaning the privacy of data should be respected. The second is Law Number 39 of 1999 concerning Human Rights, which attests that being an object of research involves asking a person for comments, opinions, or information about their personal life and data while recording their pictures and sounds [28]. This means that personal data protection is a fundamental right, and entitlement of all Indonesian people under human rights. The focus of human rights is existence and dignity, ensuring people are not trampled upon because the legal responsibility or personality of a human being is enforced above this dignity, enabling citizens to enjoy their rights and adhere to various obligations.

The protection of human rights was alluded to in the Qur'an by the word of Allah in surah al-Isra: 70, which reads: "*And indeed We have honored the children of Adam, We raised them on land and at sea, We gave them sustenance from the good things, and We gave them more advantages, perfect over most of the creatures we have created.*"

Although this verse indicates that Allah elevates the human status, many violations occur that require the intervention of one group in another. The core problem in human rights is preserving one's rights from threats, disturbances, obstacles, and challenges from other parties or damage caused by outsiders. John Locke termed this phenomenon natural rights, which should not be eliminated by any institutions and organizations, including the state, because they existed before its formation [30]. Human rights and *Maqasid Sharia* are significantly related because both are aimed at guaranteeing the benefit of humans. *Maqasid Sharia* provides an alternative to evade the abyss of difficulties when faced with urgent problems, forces, and challenging circumstances, ensuring the rights of humans are protected from damage [29].

2.4 State Governance

In the category of state governance, there were three laws relating to personal data protection. They are Law Number 18 of 2003 concerning Advocacy, which explains that the obligation to maintain confidentiality covers present and former client secrets, whose information must be kept confidential [30]. Law Number 43 of 2009 concerning Archives indicates that the civil rights of the people include social, economic, and political rights, as evidenced in archives, such as land certificates, diplomas, marriage certificates, birth certificates, resident cards, population data, wills, and business licenses [31]. There are two terms in Population Administrative Law Number 24 of 2013, namely population and personal data. Individual and structured aggregate data resulting from Population and Civil Registration activities are examples of population data, while personal data refers to specific personal information stored, maintained, kept, and protected by confidentiality [32]. In Law Number 14 of 2008 concerning Public Information Disclosure, public information may be disclosed to others when it endangers life or jeopardizes national security [33]. According to Law Number 17 of 2011 concerning State Intelligence, secret state information is sensitive information that could jeopardize the safety of the state, highlighting the criticality surrounding its leakage [34]. Finally, Judicial Commission Law Number 18 of 2011 requires judicial commission members to keep any information obtained a secret due to the nature of the information and based on the member's position [35].

These laws aim to protect the confidentiality of citizen's data, thereby obeying *Maqasid Sharia* in maintaining religion, life, and mental wellbeing. Allah gave reassurance on good governance in surah Al-Hajj: 41 that "*They are˺ those who, if established in the land by Us, would perform the prayer, pay alms-tax, encourage what is good, and forbid what is evil. Furthermore, with Allah rests the outcome of* *all affairs."* From the verse above, good governance in the context of Islamic law entails the use of authority to manage development with the goal of (1) creating a conducive environment for the community to fulfill their physical and spiritual needs, as symbolized by the enforcement of prayer, (2) initiating zakat for the creation of prosperity and welfare, and (3) establishing political stability, as inspired by *amar ma'ruf and nahi munkar* (uphold the truth and forbid what is wrong). Consequently, there are three types of governance in Islam, according to the verse, namely (a) spiritual governance, (b) economic governance, and (c) political governance [36].

2.5 Crime Prevention

Five laws on personal data protection belong to the crime prevention category. They include Law Number 31 of 1999 concerning the Corruption Criminal Act [37] and Law Number 30 of 2002 concerning the Corruption Eradication Commission [38], which have a clear objective of protecting witnesses and society's wealth from corrupt government officials. Law Number 15 of 2003 concerning Anti-Terrorism indicates that the state is responsible for the safety and security of its citizens and must ensure the rights of terrorist crime victims [39]. The provisions of Law Number 35 of 2009 concerning Drugs aim to protect witnesses, reporters, investigators, public prosecutors, and judges, as well as their families, who investigate narcotic crime cases [40]. Finally, Law Number 8 of 2010 concerning the Money Laundering Criminal Act ensures whistleblowers and witnesses in money laundering cases must be given special protection before, during, and after the case investigation process [41].

The laws above corroborate *Maqasid Sharia* in protecting witnesses and the society's wealth from any crime. This is supported by Allah's word in Surah Al-Baqarah 188: *"And do not eat the wealth among yourselves in a false way, and (do not) bribe the judges with it, with the intention that you may eat up some of the wealth of others by way of sin, even though you know."* **T**his verse is a confirmation that anyone can obtain wealth vainly. Moreover, properties received are still considered illegal even after a judge's decision because the information was misleading to ensure one's entitlement [42].

In Ministry Regulation Number 20 of 2016 concerning Personal Data Protection on Electronic Systems, article 2 point 2 (a, b, and h) only regulates the principles of Good Data Protection, including respect for privacy and confidentiality based on legislative provisions, and states that the data user is responsible for his data [13]. Article 2 of Ministry Regulation Number 20 of 2016 concerning Personal Data Protection on Electronic Systems was intended to fulfill *Maqasid Sharia* by shielding religion from persons who discriminate against Muslims as well as protect human life and sanity from physical and mental damage. Unfortunately, this regulation only provided administrative sanctions and failed to stipulate any criminal sanctions for violators. In e-commerce transactions, a person's data must be protected because it contains their profile, contact history, location, pictures, documents, and other private matters. The Qur'an emphasizes the primacy of privacy as the word of Allah in surah

An-Nur: 27, which reads: "*O you who believe, do not enter a house that is not yours before asking permission and greeting its inhabitants. That is better for you so that you (always) remember.*"

Although the Qur'an does not explain in detail how to protect personal data during e-commerce transactions, greeting and asking permission before entering a person's house means that God, through his word in surah An-Nur, protects or limits the socialization of believers [11]. It is like regulations of personal data protection, where information can only be accessed by obtaining permission from the concerned party. This supports the words of the Prophet Muhammad in a hadith quoted from the hadith *Sahih Bukhari*, "*if someone peeks into your house when you do not allow it, then you throw a stone at him so that it blinds his eyes, you will not have sinned for it.*" [43].

Personal data protection arises because of concerns about breaches that individuals and legal entities may experience, which lead to material and moral losses. The basis of norms and implementation in the Personal Data Protection Bill is based on the principles of protection, legal certainty, public interest, expediency, prudence, balance, and responsibility. Hence, this bill adheres to the objectives of *Maqasid Sharia* to protect the soul, mind, offspring, and property and to achieve justice for all. Rasulullah stated that a deviant of Allah is someone who continues to commit heinous and evil deeds even while praying. The reason for this is that Ibn Kathir stated that three things in prayer encourage a person always to do good. The three things in question are sincerity, solemnity, and remembrance of Allah [44].

3 Conclusion

This normative research on the protection of personal data on Indonesian e-commerce platforms from the view of *Maqasid Sharia* showed the general state of the law regarding this subject. This is because the rules of personal data protection are contained in several different laws and regulations and only describe the general concept. A total of 30 laws governs personal data protection in Indonesia and can be classified into several categories, namely health, finance and business, human rights, state governance, and crime prevention. The laws in all categories are in line with *Maqasid Sharia.* The health laws aim to maintain the life and lineage of patients and health workers, while the finance and business categories guard the customer's wealth, dignity, and lineage. The laws in the human rights and state governance categories intend to protect religion, mind, life, property, and lineage. Finally, the crime prevention regulations protect life, mind, offspring, and property. The categories above show that the existing personal data protection regulations in Indonesia contain the five principles of *Maqasid Sharia*, namely maintaining religion, mind, life, property, and lineage.

References

1. Dinisari MC (2020) E-commerce Dorong Perekonomian Indonesia, selama Pandemi Covid-19
2. Laming S (2020) Tren E-Commerce Pada Era Pandemi COVID-19. J Penelit Hum 11(2):55–63
3. Ustadiyanto R (2001) E-commerce framework. Yogyakarta, Andi
4. Indonesia C (2020) Cerita Lengkap Bocornya 91 Juta Data Akun Tokopedia
5. CNN Indonesia (2020) 13 Million Data Leaked by Bukalapak for Sale at Hacker Forum
6. Jawa Pos (2021) 91 Million Tokopedia account data leaked and distributed in internet forums
7. Mamduh N (2020) Lazada Confirms 1.1 million RedMart accounts leaked by hackers
8. Djafar W, Sumigar BL, Fritz BR (2016) Setianti, Perlindungan Data Pribadi—Usulan Pelembagaan Kebijakan Dari Perspektif Hak Asasi Manusia. Jakarta, Elsam
9. Rosadi SD (2018) Protecting privacy on personal data in digital economic era: legal framework in Indonesia. Brawijaya Law J 5(1):123–157
10. Sinaga EMC (2020) Formulasi Legislasi Perlindungan Data Pribadi dalam revolusi Industri 4.0. Rechtsvinding J 9(2):237–256
11. Angriani P (2021) Perlindungan Hukum terhadap Data Pribadi dalam Transaksi E-Commerce: Perspektif Hukum Islam dan Hukum Positif. J Syariah dan Huk 19(2):154
12. Gavison R (1980) Privacy and the limits of law. Yale Law J 89:421–471
13. Ministry Regulation No.20 of 2016 concerning Personal Data Protection in Electronics System (2016) pp Article 3
14. Faqih M (1994) Epistemologi Syari'ah: Mencari Format Baru Fiqh Indonesia. Walisongo Press, Semarang
15. Al-Syātibi AI (2018) Al-Muwafaqat Fi Usul Al-Shari'ah. Lebanon, Dārul kutub al-Ilmiyah
16. Ichsan M (2015) Pengantar Hukum Islam. Laboratorium Hukum, Fakultas Hukum, Universitas Muhammadiyah Yogyakarta, Yogyakarta
17. Khatib S (2018) Konsep Maqashid Al-Syari'ah: Perbandingan Antara Pemikiran Al-Ghazali dan Al-Syabiti. MIZANI Wacana Huk. 5(1):123–157
18. Law Number 29 Year 2004 cocerning The Medical Practice. Jakarta (2004) pp article 51 (c) and article 52 (e)
19. Law Number 36 of 2009 concerning Health. Indonesia (2009) pp article 8 and article 44 point 3
20. Law Number 40 of 2009 concerning Hospital (2009) pp Article 32
21. Law Number 36 of 2014 concerning Health Workers (2009) pp article 58 point 1 (c)
22. Law Number 18 of 2014 concerning Mental health. Indonesia (2014) pp article 70 point 1
23. The Republic Of Indonesia, Law Number 10 of 1998 concerning Amendments to Law Number 7 of 1992 concerning Banking article 40, Law No 21 of 2008 concerning Sharia Banking article 41, and Law No.6 of 2009 concerning the amendment of Law Number 23 of 1999 concerning Bank Indonesia. 2009, p. article 14 and 30
24. Law Number 8 of 1997 concerning Company Document. Indonesia (1997) pp article 5 and 20
25. Law Number 36 of 1999 concerning Telecommunications. Indonesia (1999) pp article 42
26. Law Number 19 of 2016 concerning Amendments to Law Number 11 of 2008 concerning Information and Electronic Transactions. Indonesia (2016) p. article 26 point 1
27. Agustin H (2021) Teori Bank Syariah. J Perbank Syariah 2(1):80
28. 1945 Constitution of the Republic of Indonesia. Indonesia (1945) p. article 21
29. Kasdi A (2014) Maqashid Syari'ah dan Hak Asasi Manusia (Implementasi Ham Dalam Pemikiran Islam). J Penelit 8(2):259
30. Law Number 18 of 2003 concerning Advocate (2003) p. article 19
31. Law Number 43 of 2009 concerning Archives. Indonesia (2009) p. article 3 and 44
32. Law Number 24 of 2013 concerning Amendments to Law Number 23 of 2006 concerning Population Administration. Indonesia (2013) p. article 84
33. Law Number 14 of 2008 concerning Public Information Disclosure. Indonesia (2003) p. article 17
34. Law Number 17 of 2011 concerning State Intelligence. Indonesia (2011) p. article 25

35. Law Number 18 of 2011 concerning Judicial Commission (2011) p. 20 A point 1 (c)
36. Setyono J (2015) Good Governance Dalam Perspektif Islam (Pendekatan Ushul Fikih: Teori Pertingkatan Norma). UIN Sunan Kalijaga Yogyakarta 6(1):36
37. Law Number 31 of 1999 concerning Corruption Criminal act. Indonesia (1999) p. Article 41
38. Law Number 30 of 2002 Corruption Eradication Comission (2002) p. Article 15
39. Law Number 15 of 2003 concerning Anti-Terrorism. Indonesia (2003) p. article 33–34(b)
40. Law Number 35 of 2009 concerning Drugs. Indonesia (2009) p. article 100 point 1 and article 106 (e)
41. Law Number 8 of 2010 concerning Money Laundering Criminal Act. Indonesia (2010) p. article 85 point 1
42. Sakinah (2014) Korupsi Dalam Perspektif Hukum Islam. Et-Tijarie 1(1):69
43. Nashirudin Al-AM (20030) *Ringkasan Shahih Bukhari*, 7 th. Gema Insani
44. Indina HR (2021) Surah Al Ankabut Ayat 45 Tentang Satu Amalan Pencegah Perbuatan Keji

Pivotal Factors Affecting Citizens in Using Smart Government Services in Indonesia

Ulung Pribadi, Juhari, Muhammad Amien Ibrahim, and Cahyadi Kurniawan

Abstract The aim of this study is to examine the analysis factors affecting citizens in using smart government services in Indonesia. A questionnaire survey was sent to 300 people who used smart government services. The collected data was analyzed using SEM-PLS. This study found that accountability, user satisfaction, trust in government, trust in technology, and perceived cost positively and significantly affected citizens to use the smart government services, with their respective *p*-values 0.032, 0.004, 0.026, and 0.044. Meanwhile, perceived risk and community culture did not significantly affect the citizens, with p-values of 0.080 and 0.170, respectively. This study only examined these affecting pivotal factors on citizens in three local districts. The next research should observe wide areas. The findings of this study can help local government stakeholders which implement smart government services, particularly in improving the perceived cost of public services. This research can strengthen user satisfaction, service innovation, and low costs so that people's trust increases in using smart government services. This study contributes to the development of the literature regarding smart government services.

Keywords E-government · Artificial intelligence · Smart government services · Local government · Public organizations

U. Pribadi (✉) · Juhari
Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: ulungpribadi@umy.ac.id

M. A. Ibrahim
Bina Nusantara University, Jakarta, Indonesia

C. Kurniawan
Government Science, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_89

1 Introduction

Currently, many governments in the world have implemented smart government services, which are sophisticated information and communication technologies to provide quality of public services to their citizens, businesses, and public agencies. Smart government in some cases uses the Internet of things (IoT) and **artificial intelligence** (AI) [1, 2]. For example, Mexico implemented applications on mobile devices for tax payments based on smart government technology [3]. Another example is that the smart government is the basis for developing Birjand city into a smarter city [4]. And the smart government has a positive impact on employee performance in the UAE [5].

Existing studies examined smart government services focusing on the technological perspective. Some examples were studies on the structure of ICT [6], software systems [7], aspects of technical innovation [8], cloud platforms [3, 9], Internet of things (IoT) [1, 10], crowdsourcing framework [11], enterprise architecture framework [12], and blockchain technology [13]. Other studies on smart government focused on government decision-making and regulation [14]. Furthermore, studies on smart government used an organizational and management perspective. This includes the transformation of planning and policy, leadership and public managers, human resources, organizational structures, bureaucracy, and budgets [15–18].

Research on smart government from the perspective of the user community is still scarce, although it is important to know what pivotal factors influence people to use smart government services. Some literature used the customers' perspective, but they only looked at the innovation and security side that customers felt [19] and the citizen-centric approach [20].

This study fills the knowledge gap by examining the pivotal factors influencing citizens in using smart government services. This study is inspired by the technology acceptance model (TAM) and then expands it with variables that are currently suitable for the conditions of local government and Indonesian society.

2 Literature Review and Theoretical Framework

2.1 Smart Government in Public Services

Smart government as a follow-up to e-government is a sophisticated technology that is applied in government to increase the government efficiency in activities and public services. Public services include services in the fields of education, population, transportation, registration, health, licensing, and others [21, 22].

The smart government uses modern ICTs to create an inter-organizational network between governmental agencies. The smart government applies smart tools and smart technology to realize the characteristics of smartness. The smart government uses social media, high devices, and mobile applications [23], with a digitalization approach [24], use the Internet of things (IoT) [25], and apply big data [26]. Smart government, implemented in the central government and local government levels, can also increase transparency and open government on data, documents, and information needed by the community, embed cooperation, collaboration, and coordination between government agencies, and provide space for interaction between the government and the community in the formulation, implementation, and evaluation of public policies, as well as fostering community engagement in the governance process [19, 20, 27–30].

2.2 *Extended Technology Acceptance Model*

Many scholars have used the technology acceptance model (TAM) models to study the behavior of people in adopting information and communication technology in government [31, 32]. Some of them used this theory to predict citizens' using of e-complaint service [33, 34], to evaluate smart government service adoption [35], to examine the use of smart government by employees [36], and to predict public value creation [37].

There are scholars who add to the theory with constructs of information quality, system quality, trust, and cost [38]. Other scholars complement the construct of perceived usefulness with the construct of user satisfaction to see the user's adoption of smart government [39]. Subsequent scholars include constructs of service quality, system quality, and information quality [40]. Another scholar complements TAM with new constructs of trust in smart government services, satisfaction, social influence, and citizen engagement [41].

2.3 *Use of Smart Government Services*

Smart government services also increase efficiency, innovation, effectiveness, openness, citizen engagement, equality, integration, creativity, sustainability, and citizen centricity [42–44]. Smart government services can provide multi-directional and personalized public services in future, transcending time, space, and region, as well as people's lifestyle [45]. Smart government services adoption necessitates 3 major stages: static, interaction, and transaction [46].

2.4 *Perceived Risk*

When the Jordanian government implemented smart government services, it paid attention to public issues which include perceived risk, perceived trust, and perceived quality [8]. Utilization of e-government services must pay close attention to the variables of trust, risk, and security [47].

2.5 Government Accountability

Government accountability means that all government activities are in accordance with the needs and interests of the wider community [48]. Scholars state the link between accountability and the smart cities development and the use of new technologies [49]. Accountability, transparency, and credibility are closely related to the use of ICTs represented by e-platforms [50].

2.6 Community Culture

Community culture is an expression in people's everyday lives informally which includes symbols, perceptions, behavior, and creation of works, including inside of organization [17]. Local cultural values have an influence on public services in the local government [51]. Scholars explain that there is a link between culture and behavior using technology in the development of the smart government [48]. Moral values (virtues, principles, and duties) and the interests and needs of the public influence the use of ICTs in the implementation of smart government field [52].

2.7 User Satisfaction

In the case of education in China, user satisfaction can increase the use of online education platforms [53]. A study found that user satisfaction is one of the indicators for the M-government adoption [54]. There is a relationship between satisfaction with adoption and public trust in the use of smart government [41]. User satisfaction is related to service quality and performance in the use of technology in smart government in the UAE [55].

2.8 Trust in Government

Residents and public servants' intentions to use smart city services in a mid-sized U.S. city are influenced by trust in the government [56]. Trust in the government and the government's Website were significant predictors of e-government service use [57]. The perception of trust and security among the millennial generation influences their use of the Internet for e-government services [47]. Smart government adoption is influenced by trust in the government [58].

2.9 Trust in Technology

Trust in technology is a person's belief that the operation of a technology can be trusted to get online information [59]. Trust in technology influences residents' and public servants' intention in using smart city services in a mid-size U.S. city [56]. Citizens' trust on the Website is a determining factor in their intention to use smart government services [57]. Trust influences citizens to adopt smart mobile government services in Jordan [60].

2.10 Perceived Costs

Many scholars considered perceived cost as an essential factor influencing user's behavioral intention to use information and communication technology [61]. Perceived cost is one of the determining factors for public employees to use technology to improve performance [62]. Cost–benefit considerations are one of the determinants that sometimes hinder the adoption of smart government [63].

3 Proposed Research Model

See Fig. 1.

4 Research Method

4.1 Data Collection

A quantitative research model, with survey techniques, has been used in this research. This study uses questionnaires as a tool to obtain primary data. This study conducted a survey of people who use smart government technology in local governments in Indonesia.

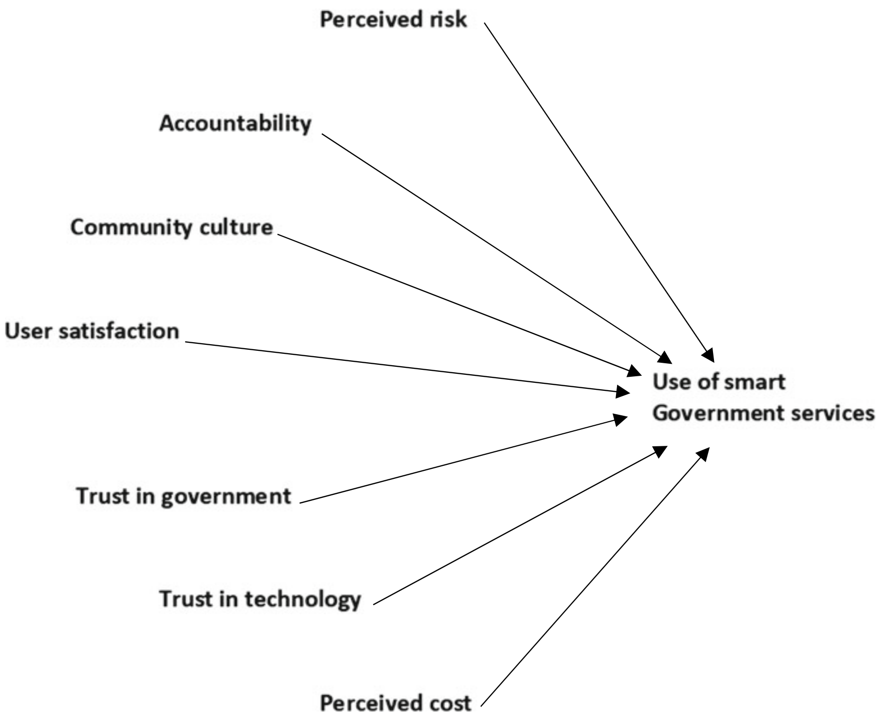


Fig. 1 Proposed research model

4.2 Sampling Technique

This study used a simple random sampling technique to select research respondents. In this context, a simple random sample was a subset of a statistical population in which each citizen who used the local government application software had the same chance of being chosen. This study uses google.form as a tool for distributing questionnaires. Respondents filled out the questions in google.form and sent them back to the researcher.

This research takes in three regions in Indonesia which represent West Indonesia, Central Indonesia, and East Indonesia. This study took 100 respondents for each region. The total respondents in this study were 300 residents who used smart government services from the local government. The calculation indicates that 300 respondents are appropriate, with a 95% confidence level and a 5% margin of error.

4.3 Measurement and Analysis Technique

To gather data, quantitative survey questions were utilized. Utilizing a Likert scale, the questions were developed. Likert scales with 1 denoting severely disagree, 2 denoting disagree, 3 denoting somewhat agree, 4 denoting agree, and 5 denoting definitely agree were used to evaluate the respondents' opinions. The data was examined using SEM-PLS to assess its reliability and validity as well as to test the validity of the hypothesis and regression.

5 Data Finding

5.1 Validated Research Model

The validity of the indicators that have been established as questionnaires is shown in Fig. 2. An indicator was regarded as valid if its value was larger than 0.5. Figure 2 demonstrates that every number was more than 0.5, demonstrating the validity of each indicator.

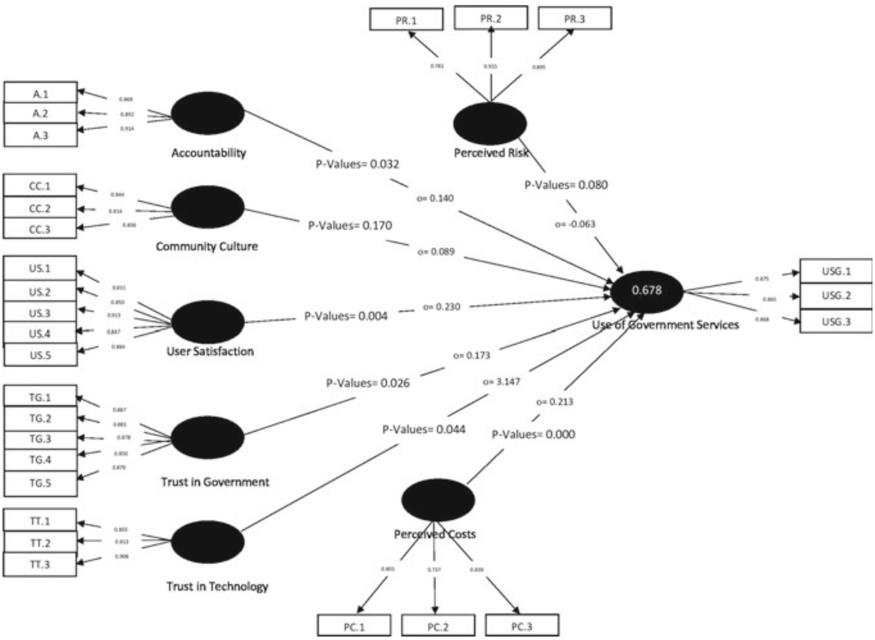


Fig. 2 Validated research model

Figure 2 depicts the results of hypothesis testing. The hypothesis is supported when the P-value is less than 0.05. The *H1* hypothesis, which stated that perceived risk positively and significantly influences the citizens using smart government services, was rejected (*p-value* = 0.080). The *H2* hypothesis, which stated that government accountability positively and significantly influences the citizens in using smart government services, was supported (*p-value* = 0.032). The *H3* hypothesis, which assumed that community culture positively and significantly influences the citizens using smart government services, was rejected (*p-value* = 0.170). The *H4* hypothesis, which stated that user satisfaction positively and significantly influences the citizens using smart government services, was supported (*p-value* = 0.004). The *H5* hypothesis, which stated that trust in government positively and significantly influences the citizens using smart government services, was supported (*p-value* = 0.026). The *H6* hypothesis, which stated that trust in technology influences citizens' use of smart government services positively and significantly, was supported (*p-value* = 0.044). The *H*7 hypothesis was supported, which stated that perceived cost influences the use of smart government services positively and significantly (*p-value* = 0.000).

6 Discussion

The study discovered that the more the government does for the benefit of the people, the more likely citizens are to use smart government services. This finding justifies the previous scholar's studies regarding the interests and needs of the wider community which can encourage citizens to trust the use of technology in government [48–50]. Moreover, this study finds that the more people feel satisfied, the more citizens are interested in using the new tool of public services. The result of this study confirms the results of earlier investigations [41, 53–55]. Furthermore, this study finds that the higher the citizens trust to the public bureaucracy, the more citizens tend to use the newest technology in the government. This finding justifies that the earlier studies regarding trust in government have proven to strengthen citizens to use the tools provided by the government [47, 56–58]. This study finds that the more citizens trust sophisticated tools, the more citizens tend to use the most recent technology in government. It supports the previous studies stated that trust in technology influences citizens' intention in using smart services [56, 57, 59, 60]. This study uncovers that the lower the cost and the shorter the time, the more citizens tend to use the most recent technology in government. It legitimized scholars' statement that perceived cost influenced citizens to use the government's technologies [61–63].

7 Conclusion

The theoretical reflection that can be built from the findings of this study is as follows: accountable government policies and programs, community satisfaction in obtaining public services, people's trust in the public bureaucracy, and perceived cost trigger citizens to use the smart government services. The practical implication suggestion is that local government must determine service fees rationally, not expensive, and affordable for the poor. This study has some limitations, including a small number of research regions (only three regencies and cities), which may not be extrapolated to the entire Indonesian territory, and a small number of respondents, which may not reflect the true situation of citizens. Future research should involve a diverse range of local government agencies in order to cover a larger geographic area, and the next studies should include a larger sample size to assess how consistent people's perceptions are.

References

1. Wirtz BW, Weyerer JC, Schichtel FT (2019) An integrative public IoT framework for smart government. Gov Inf Q 36(2):333–345. https://doi.org/10.1016/j.giq.2018.07.001
2. Kankanhalli A, Charalabidis Y, Mellouli S (2019) IoT and AI for smart government: a research agenda. Gov Inf Q 36(2):304–309. https://doi.org/10.1016/j.giq.2019.02.003
3. Cedillo-Elias EJ, Larios VM, Orizaga-Trejo JA, Lomas-Moreno CE, Ramirez JRB, Maciel R (2019) A cloud platform for smart government services, using SDN networks: the case of study at Jalisco State in Mexico. In: 2019 IEEE international smart cities conference (ISC2), Casablanca, Morocco, Oct 2019, pp 372–377. https://doi.org/10.1109/ISC246665.2019.9071680
4. Ghasemi A, Saberi M (2020) The key factors in transforming Birjand city to a smart city: smart mobility, smart government. Indones J Electr Eng Comput Sci 19(1):317–324. https://doi.org/10.11591/ijeecs.v19.i1.pp317-324
5. Alfalasi K, Ameen A, Isaac O, Khalifa GSA, Midhunchakkaravarthy D (2020) Impact of actual usage of smart government on the net benefits (knowledge acquisition, communication quality, competence, productivity, decision quality). TEST Eng Manage 82:14770–14782
6. Scholl HJ AlAwadhi S (2016) Creating smart governance: the key to radical ICT overhaul at the city of Munich. Inf Polity 21(1):21–42. https://doi.org/10.3233/IP-150369
7. Fajar AN, Nugeraha Utama D (2018) SGSC framework: smart government in supply chain based on FODA. Bull Electr Eng Inf 7(3):411–416. https://doi.org/10.11591/eei.v7i3.817
8. Jaradat M-IRM, Moustafa AA, Al-Mashaqba AM (2018) Exploring perceived risk, perceived trust, perceived quality and the innovative characteristics in the adoption of smart government services in Jordan. Int J Mob Commun 16(4):399–439
9. Witanto JN, Lim H, Atiquzzaman M (2018) Smart government framework with geo-crowdsourcing and social media analysis. Future Gener Comput Syst 89:1–9. https://doi.org/10.1016/j.future.2018.06.019
10. Chatfield AT, Reddick CG (2019) A framework for Internet of Things-enabled smart government: a case of IoT cybersecurity policies and use cases in U.S. federal government. Gov Inf Q 36(2):346–357. https://doi.org/10.1016/j.giq.2018.09.007
11. Puritat K (2019) A gamified mobile-based approach with web monitoring for a crowdsourcing framework designed for urban problems related smart government: a case study of Chiang Mai,

Thailand. Int J Interact Mob Technol IJIM 13(12):55–66. https://doi.org/10.3991/ijim.v13i12.10989

12. Cherrabi M, Benbrahim M, Boutahar J (2020) Adaptive enterprise architecture M-NEA for Moroccan national system: towards Moroccan smart-government. In: 2020 IEEE international conference of Moroccan geomatics (Morgeo), Casablanca, Morocco, May 2020, pp 1–8. https://doi.org/10.1109/Morgeo49228.2020.9121896
13. Shan S, Duan X, Zhang Y, Zhang TT, Li H (2021) Research on collaborative governance of smart government based on blockchain technology: an evolutionary approach. Discrete Dyn Nat Soc 2021:1–23. https://doi.org/10.1155/2021/6634386
14. Kennedy R (2016) E-regulation and the rule of law: smart government, institutional information infrastructures, and fundamental values. Inf Polity 21(1):77–98. https://doi.org/10.3233/IP-150368
15. Kravchenko AG, Litvinova SF (2015) The prospects for legislative modeling 'Smart Government' in political and legal realities of Russia. Mediterr J Soc Sci 6(3):341–346. https://doi.org/10.5901/mjss.2015.v6n3p341
16. Al-Obthani F, Ameen A (2019) Association between transformational leadership and smart government among employees in UAE public organizations. Int J Emerg Technol 10(1a): 98–104
17. Melati C, Janissek-Muniz R (2020) Smart government: analysis of dimensions from the perspective of public managers. Rev Adm Pública 54(3):400–415. https://doi.org/10.1590/0034-761220190226x
18. Sensuse DI, Arief A, Mursanto P (2022) An empirical validation of foundation models for smart government in Indonesia. Int J Adv Sci Eng Inf. Technol. 12(3):1132. https://doi.org/10.18517/ijaseit.12.3.13442
19. Hashim KF, Hashim NL, Ismail S, Miniaoui S, Atalla S (2020) Citizen readiness to adopt the new emerging technologies in Dubai smart government services. In: 2020 6th international conference on science in information technology (ICSITech), Palu, Indonesia, Oct 2020, pp 1–5. https://doi.org/10.1109/ICSITech49800.2020.9392071
20. Obedait AA, Youssef M, Ljepava N (2019) Citizen-centric approach in delivery of smart government services. In: Al-Masri A, Curran K (eds) Smart technologies and innovation for a sustainable future. Springer, Cham, pp 73–80. https://doi.org/10.1007/978-3-030-01659-3_10
21. Sanjifa ZN, Sumpeno S, Suprapto YK (2019) Community feedback analysis using latent semantic analysis (LSA) to support smart government. In: 2019 International seminar on intelligent technology and its applications (ISITIA), Surabaya, Indonesia, Aug 2019, pp 428–433. https://doi.org/10.1109/ISITIA.2019.8937137
22. Hermanto A, Binti Ibrahim R, Kusnanto G (2020) Improving value-based e-government towards the achievement of smart government. In: 2020 Fifth international conference on informatics and computing (ICIC), Gorontalo, Indonesia, Nov 2020, pp 1–7. https://doi.org/10.1109/ICIC50835.2020.9288609
23. Algebri HK, Husin Z, Abdulhussin AM, Yaakob N (2017) Why move toward the smart government. In: 2017 international symposium on computer science and intelligent controls (ISCSIC), Budapest, Oct 2017, pp 167–171. https://doi.org/10.1109/ISCSIC.2017.34
24. Sankowska P (2018) Smart government: an European approach toward building sustainable and secure cities of tomorrow. Int J Technol 9(7):1355. https://doi.org/10.14716/ijtech.v9i7.2517
25. Al Enezi A, Al Meraj Z, Manuel P (2018) Challenges of IoT based smart-government development. In: 2018 IEEE green technologies conference (GreenTech), Austin, TX, Apr 2018, pp 155–160. https://doi.org/10.1109/GreenTech.2018.00036
26. Zhang S, Lan Y (2019) Study on smart government construction of big data-oriented. J Phys Conf Ser 1288(1):012073. https://doi.org/10.1088/1742-6596/1288/1/012073
27. Gil-Garcia JR, Helbig N, Ojo A (2014) Being smart: emerging technologies and innovation in the public sector. Gov Inf Q 31:I1–I8. https://doi.org/10.1016/j.giq.2014.09.001
28. Sawafi AMA Awad MA (2020) Citizen engagement in smart government: content analysis of Mohammed Bin Rashid tweets. In:2020 14th international conference on innovations in

information technology (IIT), Al Ain, United Arab Emirates, Nov 2020, pp 160–164. https://doi.org/10.1109/IIT50501.2020.9299046
29. Fu'adi DK, Arief A, Sensuse DI, Syahrizal A (2020) Conceptualizing smart government implementation in smart city context: a systematic review. In: 2020 fifth international conference on informatics and Computing (ICIC), Gorontalo, Indonesia, Nov 2020, pp 1–7. https://doi.org/10.1109/ICIC50835.2020.9288656
30. Vujković P, Ravšelj D, Umek L, Aristovnik A (2022) Bibliometric analysis of smart public governance research: smart city and smart government in comparative perspective. Soc Sci 11(7):293. https://doi.org/10.3390/socsci11070293
31. Susanto TD, Diani MM, Hafidz I (2017) User acceptance of e-government citizen report system (a case study of City113 app). Procedia Comput. Sci. 124:560–568. https://doi.org/10.1016/j.procs.2017.12.190
32. Adiyarta K, Napitupulu D, Nurdianto H, Rahim R, Ahmar A (2018) User acceptance of e-government services based on TRAM model. IOP Conf Ser Mater Sci Eng 352:012057. https://doi.org/10.1088/1757-899X/352/1/012057
33. Alryalat MAA (2017) Measuring citizens' adoption of electronic complaint service (ECS) in Jordan: validation of the extended technology acceptance model (TAM). Int J Electron Gov Res 13(2):47–65. https://doi.org/10.4018/IJEGR.2017040103
34. Pribadi U (2021) Citizens' intention to use e-government services: the case of e-complaint service in Indonesia. Int J Electron Gov 13(2):114–131. https://doi.org/10.1504/IJEG.2021.116884
35. Mensah IK (2018) E-government services adoption: the important elements of trust and transparency. Int J Electron Gov Res 14(3):12–31. https://doi.org/10.4018/IJEGR.2018070102
36. Ameen A, Alfalasi K, Gazem NA, Isaac O (2019) Impact of system quality, information quality, and service quality on actual usage of smart government. In: 2019 first international conference of intelligent computing and engineering (ICOICE), Hadhramout, Yemen, Dec 2019, pp 1–6. https://doi.org/10.1109/ICOICE48418.2019.9035144
37. Chohan SR, Hu G (2020) Success factors influencing citizens' adoption of IoT service orchestration for public value creation in smart government. IEEE Access 8:208427–208448. https://doi.org/10.1109/ACCESS.2020.3036054
38. Weerakkody V, Irani Z, Lee H, Hindi N, Osman I (2016) Are U.K. citizens satisfied with e-government services? Identifying and testing antecedents of satisfaction. Inf Syst Manag 33(4):331–343. https://doi.org/10.1080/10580530.2016.1220216
39. Kurfalı M, Arifoğlu A, Tokdemir G, Paçin Y (2017) Adoption of e-government services in Turkey. Comput Hum Behav 66:168–178. https://doi.org/10.1016/j.chb.2016.09.041
40. Al-Obthani F, Ameen A (2019) Influence of overall quality and innovativeness on actual usage of smart government: an empirical study on the UAE public sector. Int J Emerg Technol 10(1a):141–146
41. Hartanti FT, Abawajy JH, Chowdhury M, Shalannanda W (2021) Citizens' trust measurement in smart government services. IEEE Access 9:150663–150676. https://doi.org/10.1109/ACCESS.2021.3124206
42. Gil-Garcia JR, Zhang J, Puron-Cid G (2016) Conceptualizing smartness in government: An integrative and multi-dimensional view. Gov Inf Q 33(3):524–534. https://doi.org/10.1016/j.giq.2016.03.002
43. Alghawi K, Ameen A, Bhaumik A (2019) Empirical study of the UAE-based smart government's characteristics and its effect on performance quality. Int J Emerg Technol 10(1a):59–65
44. Alghawi K, Ameen A, Bhaumik A (2019) The role of smart government characteristics for enhancing UAE's public service quality. Int J Emerg Technol 10(1a):01–07
45. Shi D, Tian Z (2020) The current situation of China's governance from the perspective of smart government. In: 2020 International conference on big data economy and information management (BDEIM), Zhengzhou, China, Dec 2020, pp 137–141. https://doi.org/10.1109/BDEIM52318.2020.00040

46. Althunibat A et al (2021) Sustainable applications of smart-government services: a model to understand smart-government adoption. Sustainability 13(6):3028. https://doi.org/10.3390/su13063028
47. Assegaff S, Andrianti A, Astri LY (2021) Evaluation of the factors influencing the trust of Millennial citizens in e-government. J Phys Conf Ser 1898(1):012009. https://doi.org/10.1088/1742-6596/1898/1/012009
48. Arief A, Sensuse DI (2018) Designing a conceptual model for smart government in Indonesia using Delphi 2nd round validity. In: 2018 International conference on advanced computer science and information systems (ICACSIS), Yogyakarta, Oct 2018, pp 93–98. https://doi.org/10.1109/ICACSIS.2018.8618239
49. Grossi G, Meijer A, Sargiacomo M (2020) A public management perspective on smart cities: 'Urban auditing' for management, governance and accountability. Public Manage Rev 22(5):633–647. https://doi.org/10.1080/14719037.2020.1733056
50. Gil O, Cortés-Cediel ME, Cantador I (2019) Citizen participation and the rise of digital media platforms in smart governance and smart cities. Int J E-Plan Res 8(1):19–34. https://doi.org/10.4018/IJEPR.2019010102
51. Pribadi U, Kim H (2021) Impacts of cultural behavior of civil servants on citizens' satisfaction: a survey on licensing services of Indonesian local government agencies. J Public Aff e2662:1–9. https://doi.org/10.1002/pa.2662
52. Yaghi A, Al-Jenaibi B (2018) Happiness, morality, rationality, and challenges in implementing smart government policy. Public Integr 20(3):284–299. https://doi.org/10.1080/10999922.2017.1364947
53. Chen T, Peng L, Yin X, Rong J, Yang J, Cong G (2020) Analysis of user satisfaction with online education platforms in China during the COVID-19 pandemic. Healthcare 8(3):200. https://doi.org/10.3390/healthcare8030200
54. Junnonyang E (2021) Integrating tam, perceived risk, trust, relative advantage, government support, social influence and user satisfaction as predictors of mobile government adoption behavior in Thailand. Int J Ebus Egov Stud 13(1):159–178. https://doi.org/10.34109/ijebeg.202113108
55. Ameen A, Al-Ali D, Isaac O, Mohammed F (2020) Examining relationship between service quality, user satisfaction, and performance impact in the context of smart government in UAE. Int J Electr Comput Eng IJECE 10(6):6026–6033. https://doi.org/10.11591/ijece.v10i6.pp6026-6033
56. Habib A, Alsmadi D, Prybutok VR (2020) Factors that determine residents' acceptance of smart city technologies. Behav Inf Technol 39(6):610–623. https://doi.org/10.1080/0144929X.2019.1693629
57. Mensah IK, Luo C, Abu-Shanab E (2021) Citizen use of e-government services websites: a proposed e-government adoption recommendation model (EGARM). Int J Electron Gov Res 17(2):19–42. https://doi.org/10.4018/IJEGR.2021040102
58. Almuraqab NAS, Jasimuddin SM, Mansoor W (2021) An empirical study of perception of the end-user on the acceptance of smart government service in the UAE. J Glob Inf Manage 29(6):1–29. https://doi.org/10.4018/JGIM.20211101.oa11
59. Kamalrudin M, Thaiban HHM, Sidek S, Hakimi H (2019) Research on trust model in online information of smart government. Int J Recent Technol Eng IJRTE 8(2S11):762–767. https://doi.org/10.35940/ijrte.B1124.0982S1119
60. Alkhwald AF, Al-Ajaleen RT (2022) Toward a conceptual model for citizens' adoption of smart mobile government services during the COVID-19 pandemic in Jordan. Inf Sci Lett 11(2):573–579. https://doi.org/10.18576/isl/110225
61. Kuo Y-F, Yen S-N (2009) Towards an understanding of the behavioral intention to use 3G mobile value-added services. Comput Hum Behav 25(1):103–110. https://doi.org/10.1016/j.chb.2008.07.007

62. Eom S-J, Choi N, Sung W (2016) The use of smart work in government: empirical analysis of Korean experiences. Gov Inf Q 33(3):562–571. https://doi.org/10.1016/j.giq.2016.01.005
63. Schedler K, Guenduez AA, Frischknecht R (2019) How smart can government be? Exploring barriers to the adoption of smart government. Inf Polity 24(1):3–20. https://doi.org/10.3233/IP-180095

Cybersecurity for Industrial IoT, Threats, Vulnerabilities, and Solutions: A Brief Review

Andrea Sánchez-Zumba and Diego Avila-Pesantez

Abstract The Industrial Internet of Things (IIoT) refers to use connected devices and technology in industrial settings such as manufacturing, energy, and transportation that links intelligent sensors and actuators. Cybersecurity in IIoT environments has become a significant issue to be solved due to the increase in attacks. The review included 43 studies conducted between 2017 and 2022 using the STRIDE model to identify and classify security threats and vulnerabilities and to develop appropriate countermeasures. The security solutions include using secure communication protocols, implementing security controls such as firewalls and intrusion detection systems, and using network segmentation and security information. It helps mitigate these risks and attacks from threats and vulnerabilities and ensures the availability, integrity, and confidentiality of the data and systems involved.

Keywords Cybersecurity · Industrial Internet of Things · STRIDE · Threats · Vulnerabilities · Systematic literature review

1 Introduction

In recent years, with the advent of the Fourth Industrial Revolution, also known as Industry 4.0, and due to the global pandemic of the SARS-CoV-2 coronavirus, the number of devices connected to the Internet grew exponentially in homes, businesses, and industries. In the manufacturing sector, IIoT technology is used to accomplish critical tasks such as automation, monitoring processes, and machine

A. Sánchez-Zumba (✉)
Pontificia Universidad Católica del Ecuador Sede Ambato (PUCESA), Ambato, Ecuador
e-mail: apsanchezz@pucesa.edu.ec; ap.sanchez@uta.edu.ec

Universidad Técnica de Ambato (UTA), Ambato, Ecuador

D. Avila-Pesantez
Grupo de Investigación en Innovación Científica y Tecnológica (GIICYT), Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador
e-mail: davila@espoch.edu.ec

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6_90

maintenance to create better business opportunities. IIoT devices include sensors, cameras, and other devices that collect data on the performance of industrial equipment and processes, which can then improve efficiency, reduce downtime, and increase overall productivity [1]. Therefore, the main objective for cybercriminals is to gain illegitimate access to restricted information to manipulate sensors, actuators, Programmable Logic Controllers (PLC), Supervisory Control and Data Acquisition Systems (SCADA), Distributed Control Systems (DCS), Industrial Control Systems (ICS) [1–3].

Cybersecurity in IIoT infrastructures has been affected by the high increase in attacks that exploit vulnerabilities in IIoT devices [4]. Undoubtedly, hackers benefit from the lack of safe industry standards and technical norms and the lack of interoperability between multi-vendor devices with low computational power, making it challenging to implement an existing security module or a unified security method [3], opening the way for attacks on the organization, massive theft of sensitive data, system manipulation, backdoors, brute force attacks [5], eavesdropping, phishing, social engineering, SQL injection, among others [3, 6–9].

The analysis of vulnerabilities, threats, risks, and security countermeasures for IIoT environments, is based on an adaptation of the Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, and Elevation of privilege (STRIDE)modeling methodology [10]. The STRIDE model is considered the most mature model and allows the division of the system into components to determine how an intruder can attack an IIoT system and how to implement defenses [11].

Based on Kitchenham's methodology [12], the main objective of this brief review is to contribute to cybersecurity in confidentiality, integrity, and availability of IIoT systems by focusing on the STRIDE threat modeling method.

2 Research Methodology

This study conducts a systematic review of the literature, through which empirical and theoretical evidence can be gathered from primary studies to answer the proposed research questions. The phases for development are Planning the Review, Conducting the Review, and Analysis.

2.1 Planning the Review

This review analyzes IIoT environments and cybersecurity mechanisms applied as a solution against threats and vulnerabilities for systems and devices involved in this complex structure. Three research questions were proposed to meet this objective:

Q1: What kind of vulnerabilities, risks, and threats exist in IIoT environments?

Q2: What are the main cyber-attacks identified in IIoT environments?

Table 1 Selection criteria

Inclusion criteria	Exclusion criteria
Articles on threats, vulnerabilities, and solutions related to IIoT cybersecurity	Website information
Articles on tools, and attack mitigation for IIoT	Articles with topics on cybersecurity in IoT, but not in IIoT
Relevant articles from primary sources related to research questions	Theses, books, posters, and editorials

Q3: What security countermeasures have been implemented to mitigate attacks in IIoT environments?

For the study, a search was conducted in the electronic databases of IEEE Xplore, Science Direct Elsevier, ACM, Springer, and MDPI, related to cybersecurity and IIoT, identifying as sources of information academic journals published between 2017 and 2022. The search strategy was based on aspects related to the research questions, using the following keywords: (1) cybersecurity, (2) IIoT, related to ("penetration tests" OR "threats" OR "attacks" OR "vulnerabilities"). In addition, to refine the selection, inclusion and exclusion criteria were applied (see Table 1).

2.2 Conduction the Review

The articles were selected in this phase considering the search strings and selection criteria. In each one, the titles, abstracts, and conclusions were reviewed, which made it possible to determine the level of contribution to each of the questions proposed. As a result of the search, 615 documents were identified, of which 43 were selected that met the established criteria (see Fig. 1).

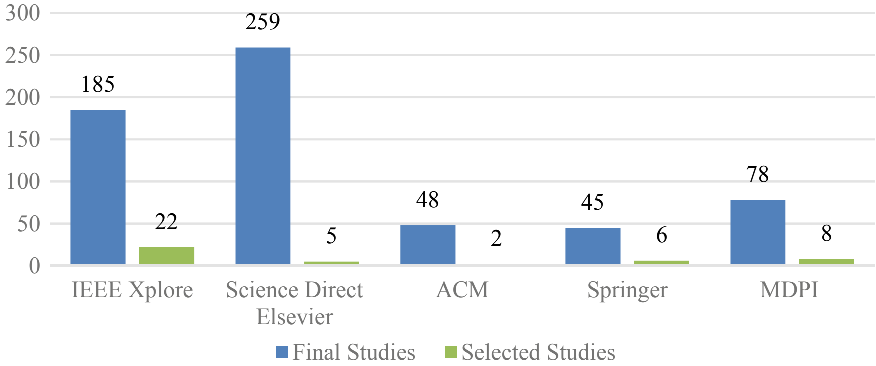


Fig. 1 Research analyzed for systematic review

2.3 Analysis

The following answering determined the inherent risks within the IIoT systems:

Q1: What kind of vulnerabilities, risks, and threats exist in IIoT environments?

There are a wide variety of risks associated with cybersecurity compromise in IIoT, with legacy devices being particularly vulnerable, with no software or firmware updates [13], combined with the use of protocols (such as Modbus) that do not incorporate encryption, authentication or authorization, allowing exploitation by malicious software or unauthorized users [14].

Spoofing An unauthorized person attempts to gain unauthorized access to steal data instead of damaging them [15, 16]. The attacker impersonates another device by using spoofed IP addresses and manipulating the set-point value of controllers [10]. In industrial environments, this kind of illegal practice is dangerous because sensors could emit wrong information, and actuators could execute actions that do not correspond to them [17, 18]. In both cases, the production chain would be seriously affected, generating economic losses and damage to industrial equipment. The risk arises when authentication between IIoT devices occurs only at the beginning of the session with unsecured tunneling protocols [19]. Neither the firmware is updated, nor the network traffic is constantly monitored. Because of these threats, data confidentiality, integrity, and authenticity could always be compromised [20].

Tampering An intruder can manipulate and send erroneous data to change a production process, change the behavior of devices and machinery to perform unsafe actions, modify quantities of elements to be produced or quantities of chemical compounds to be used depending on the area of the industrial sector without authorization in order to gain access to confidential information [18, 21, 22]. Vulnerabilities remain the obsolete technologies and industry standards without standardization and security with which IIoT devices, such as sensors and actuators, are manufactured, facilitating data interception or physical manipulation through backdoor attacks [17, 23, 24]. As a result, the integrity and availability of industrial data are not guaranteed.

Repudiation A threat in which the system cannot trace malicious or prohibited activities. This threat results from the data's lack of validation and integrity [18, 20], and the attackers may deny acting to evade accountability. This threat appears because of using an unsecured tunneling protocol.

Information Disclosure Information Leakage occurs when an attacker manages to eavesdrop on the communication between IIoT devices or when an application in the Industrial environment inadvertently discloses information to unauthorized users [25]. The intruder also attempts to read a file to which access was not granted or to read data in transit without authorization to affect processes, and data flows. At the industrial level, Confidentiality is affected, and it could represent patent theft [19, 23].

Deny of Service (DoS) Malicious actors generally use a sensor or actuator connected to the network. However, it can be a gateway, as a way to enter and flood the industrial system with bogus traffic forcing them to serve malicious requests so that they cannot complete the job they were intended for by sabotaging a process [18, 26]. The integrity and availability are compromised in IIoT systems. On the other hand, DDoS attacks are volumetric and turn system components into zombies controlled by crackers. The result is a degradation in data quality during the processing of requests.

Elevation of Privileges (EoP) They are also called escalation of privilege or privilege escalation. This threat is like spoofing, but instead of impersonating an identity, the attacker looks at privileged access to resources to gain unauthorized access to information and compromise a system with administrator permissions [6, 18, 19, 22, 23], affecting the availability, integrity, and confidentiality of the industrial process (see Table 2). It is worth noting that IIoT environments are critical infrastructure, and the consequences of a successful attack can be severe, as it can cause physical damage or loss of human lives. Therefore, it is crucial to have a robust security strategy in place to protect against these types of attacks.

Q2: Which are the main cyber-attacks identified in IIoT environments?

Ethical hackers perform cyber-attacks using the threats and vulnerabilities detailed in the previous section, combined with the evil practices of insecure passwords, unencrypted data, and uninstalled security patches. The main cyber-attacks in IIoT are shown below.

Spoofing The main cyber-attacks related to spoofing in IIoT go hand in hand with device spoofing creating fake devices to gain unauthorized access [25]. Man in the Middle (MiTM) attacks [31], with their variants Sniffing, Session Hijacking, and Packet Injection attacks, may intercept and modify communication between devices to disrupt communication [3, 9, 10]. Also include Phishing with fake emails or websites that trick users into providing private information or credentials [5, 31]. DNS spoofing redirects traffic to a fake website that appears legitimate to install malware. ARP spoofing allows fake ARP entries in a network looking to intercept

Table 2 Summary of vulnerabilities, risk, and threats IIoT

Threat modeling	Threat	Security property breached	References
STRIDE [14, 16, 17]	Spoofing	Confidentiality, integrity, authenticity	[17–19]
	Tampering	Integrity, availability	[21, 23, 24]
	Repudiation	Integrity	[23, 27]
	Information disclosure	Confidentiality	[19, 23, 25]
	Deny of service (DoS)	Integrity, availability	[26, 28–30]
	Elevation of privileges (EoP)	Confidentiality, integrity, authorization	[6, 19, 23]

traffic or perform MitM attacks [3, 32]. Finally, IP spoofing uses a fake IP address to conceal their identity and gain access to a network or launch a DDoS attack [5].

Tampering Tampering attacks such as Cross-Site Scripting (XSS) and SQL malware injection damage the integrity of the industrial system [8, 32]. It also includes MiTM attacks, intended to steal network control and can eavesdrop on the communication between devices and falsify information exchanged with malicious intent [3, 25]. Furthermore, attackers modify the firmware, change the configuration of an IIoT device and even replace it with a malicious one to change its behavior and gain unauthorized access to network, functions, or sensitive data, to disrupt its operation or cause a failure in an industrial system [32].

Repudiation It can be affected by Brute Force and Dictionary Attacks [9, 21, 35], which can exploit this threat to illegally access the industrial system and steal information without leaving traces in simple ways through email and phishing [23, 27, 33]. Command Injection is another acute attack where the cracker injects commands into an IIoT device and denies responsibility for the actions taken. These attacks could be combined with replay attacks, false alarms, and false reports.

Deny of Service (DoS) A DoS-related attack is a Replay attack, which maliciously replays traffic repeatedly to a specific destination to affect the performance of the process flooding an IIoT device with traffic to disrupt its operation or cause a failure [3, 21, 29]. Malicious Code Injection, smurfing, and ping of Death increase the amount of traffic sent to a device or network, overwhelming it and causing a failure [3, 32, 34]. Teardrop attacks attempt to make a computer resource unavailable by flooding a network or server with requests and data, and jamming attacks are other types of DoS attacks [3, 5, 8]. Attackers also use resource depletion to consume resources such as memory, storage, or processing power to disrupt the operation of a device.

Elevation of Privileges EoP attacks in the IIoT involve attackers gaining access to a system or device with higher-level permissions than they should have. Examples of EoP attacks in IIoT include Malware such as viruses or Trojans and Ramsonware [3, 5]. On the other hand, an attacker can exploit a buffer overflow vulnerability to execute code with higher privileges commonly related to default or weak credentials [32]. Social engineering can be used to trick an employee into giving up their privileged credentials [8]. Moreover, finally, misconfigured devices allow attackers to gain access (see Table 3).

Q3: What security countermeasures have been implemented to mitigate attacks in IIoT environments?

The security solutions to protect the IIoT infrastructure must be designed not to interrupt or affect operations [35] to guarantee the industrial system's confidentiality, integrity, and availability. Thus, the ways to mitigate IIoT attacks are described [10].

Spoofing Confidentiality is compromised, and the main countermeasures are authentication using One Time Password (OTP), multifactor authentication (MFA),

Table 3 Summary of cyber-attacks in IIoT environments

Cyber-attacks	References	S	T	R	I	D	E
Device spoofing	[8, 25]	x					
MiTM	[3, 10, 31]	x	x				
Sniffing	[3, 10]	x					
Session hijacking	[3, 8, 9]	x					
Phishing and social engineering	[5, 8, 31, 33]	x			x		x
DNS, ARP, and IP spoofing	[3, 5, 32]	x					
Cross-site scripting	[8, 32]		x				
SQL and packet injection	[3, 7, 8, 32]	x	x	x	x	x	
Firmware modification	[5, 32]		x				
Brute force	[5, 9]			x			
Dictionary attacks	[9]			x			
Replay attacks	[3, 21, 29]			x		x	
Side channel, eavesdropping, and teardrop	[3, 9, 25, 31]				x	x	
Jamming	[3, 5, 8]					x	
Malware	[3, 5, 33]						x
Buffer overflow	[8, 32]						x
Ping of death	[3, 34]	x				x	

ACPKC-based two level verification [7, 27, 29]. In IIoT, these measures must be applied to actuators, sensors, PLCs, and HMIs before receiving or transmitting data to ensure that the information originates from a legitimate device rather than a fraudulent source [36]. In addition, it is essential to use cryptographic algorithms with symmetric or asymmetric keys for two-way authentication, such as secure hash (SHA-x), hash-based message authenticated code (HMAC), the elliptic curve digital signature algorithm (ECDSA) for asymmetric keys [29, 37] and ACPKC-based two level verification.

Tampering Data manipulation affects the integrity of information. Some of the countermeasures are the use of hash functions (SHA-256, MDS, ERE), HMAC of TLS, applying hardware-based VPNs [3, 38], and encrypting data with strong quantum cryptography [7, 39]. These protection techniques should be used mainly on sensors and actuators within the industrial process. Also, it can be mitigated with intrusion detection and prevention systems [7, 40] and finally with authentication [7, 41] and authorization systems.

Repudiation It also affects data integrity. Nonrepudiation is sought, for which the initial countermeasure is the activation of audit logs for data access/sending and logging failed access attempts [7, 18, 27]. In addition, it is suggested to use secure communication protocols, digital signatures, implementing intrusion detection systems [27, 40], Hardware Root of Trust (HRoT) security credentials, hardware security module (HSM) [42], monitoring devices and authentication methods.

Information Disclosure The way to ensure the confidentiality of the information is by implementing Cryptography [42] and Encryption (symmetric/asymmetric mechanisms) AES-256, RSA-4096, ECC, Secure Element (HSM), Authenticated Encryption of TLS, AES, IPSec, message encryption/sign-encryption (RSA, DSA, IBE, ABE, ERE), Account locking, Delayed response, and multifactor authentication schemes [7, 22, 27]. These measures focus on taking care of the information to and from sensors, actuators, PLCs, and gateways, which are the core of an industrial process.

Deny of Services The way to avoid this threat is by implementing redundant components/networks, secure elements (HSM), data rate limiting, access control, authentication, and authorization system [27, 41]. Also, configure Multi-Level DDoS Mitigation Framework (MLDMF) [29]. Intrusion detection systems based on signatures and statistical anomalies [38, 40] raise next generation firewalls with improved traffic filtering capabilities [5, 43] combined with VPN [38].

Elevation of Privilege Implement access control systems [27], apply authorization using the least privilege principle [18], and implement service provider security policies [42]. It is suggested to implement a firewall and proxy [43]. See Table 4.

Table 4 Summary of main security countermeasures in IIoT

Countermeasures	References	S	T	R	I	D	E
OTP	[14, 22, 29]	x					
Multifactor authentication	[7, 22, 24, 27, 41]	x	x	x	x	x	x
Authorization methods	[18, 41]	x			x	x	x
Hash functions	[3, 37]		x				x
Auditing and logging	[7, 18, 27]			x			x
Encryption (AES, RSA, ECC)	[7, 22, 27, 29, 43]	x	x	x	x	x	x
Cryptography	[39, 41, 42]	x			x		
Firewall and proxy	[5, 43]		x			x	
Intrusion detection systems	[7, 38, 40, 43]	x	x			x	x
Access control systems	[27, 41, 44]			x		x	x
Security policies	[25, 42]			x			x
VPN	[3, 38]	x	x			x	

3 Discussion

According to the analysis, among the main threats detected in IIoT environments is spoofing, which according to [3], combines other attacks such as MiTM, Sniffing, Session Hijacking, Packet Injection attacks, which broaden the spectrum of the attack vector in the industry. The DoS threat that affects the integrity and availability of information, and according to [3, 18, 26], has been one of the most frequent because it uses attacks such as Replay attacks, Malicious Code Injection, flooding, smurfing, ping of death, teardrop, and jamming attacks to achieve its goal [3]. Additionally, an attack, although not very popular, EoP is a point of concern because it affects the CIA triad when the attacker gains access with administrator permissions to take complete control of the industrial intelligent control system.

Several authors agree that the main countermeasures for IIoT threats are based on multifactor authentication mechanisms, authorization methods, encryption techniques (AES, RSA, ECC), and hash functions. Various articles mention ways to mitigate the threats analyzed with STRIDE, where intrusion detection and prevention systems combined with access control systems, enabling access logs, firewall, proxy, VPN, and firmware updates in sensors, actuators, PLC, Gateway, and industrial end devices should be incorporated.

4 Conclusions

There are few security mechanisms to protect industrial networks due to the diversity of protocols and the need for unified, secure standards that complicate efforts to introduce protection to these critical systems. This insecurity is mainly due to low computational capacity, obsolete technology, and the extensive workload of sensors and actuators. However, the standards differ from IIoT, which proposes that security regulations for IIoT should be established as soon as possible due to the risks they pose to the IIoT environment.

Also, it has been evidenced that cyber-attacks on critical, intelligent factory systems have increased in the last year by approximately 50%, demonstrating the importance of IIoT security. Nevertheless, insufficient attention from top management, limited budget, and untrained human factors are the primary cybersecurity challenges that IIoT device manufacturers and vendors must overcome. In this sense, the low availability of correct tools and processes, consistent methodologies to detect threats, attacks, and vulnerabilities, and the establishment of efficient ways of mitigation, generate alarms worldwide. The use of the STRIDE threat modeling method has a positive impact and increases the credibility of the study by providing a well-established methodology that is widely recognized in the cybersecurity community. Future research could combinate the use of artificial intelligence and machine learning techniques to improve the security of IIoT systems and devices.

References

1. Chaudhary S, Gupta K, Johari R, Bhatnagar A, Bhatia R (2019) CRAIoT: concept, review and application(s) of IoT. In: 2019 4th International conference on Internet of Things: smart innovation and usages (IoT-SIU)
2. Sen S, Song L (2021) An IIoT-based networked industrial control system architecture to secure industrial applications. In: IEACon 2021–2021 IEEE industrial electronics and applications conference. Institute of Electrical and Electronics Engineers Inc., pp 280–285
3. Kim HM, Lee KH (2022) IIoT malware detection using edge computing and deep learning for cybersecurity in smart factories. Appl Sci (Switzerland) 12. https://doi.org/10.3390/app12157679
4. Nimmy K, Sankaran S, Achuthan K, Calyam P (2022) Securing remote user authentication in industrial Internet of Things. In: Proceedings—IEEE consumer communications and networking conference, CCNC. Institute of Electrical and Electronics Engineers Inc., pp 244–247
5. Tsiknas K, Taketzis D, Demertzis K, Skianis C (2021) Cyber threats to industrial IoT: a survey on attacks and countermeasures. IoT 2:163–186. https://doi.org/10.3390/iot2010009
6. Lackner M, Markl E, Aburaia M (2018) Cybersecurity management for (industrial) Internet of Things: challenges and opportunities. J Inf Technol Softw Eng 08. https://doi.org/10.4172/2165-7866.1000250
7. Khondoker R, Magin D, Bayarou K (2015) Security analysis of OpenRadio and SoftRAN with STRIDE framework
8. Chu G, Lisitsa A (2019) Penetration testing for Internet of Things and its automation. In: Proceedings—20th international conference on high performance computing and communications, 16th international conference on smart city and 4th international conference on data science and systems, HPCC/SmartCity/DSS 2018. Institute of Electrical and Electronics Engineers Inc., pp 1479–1484
9. Alanazi R, Aljuhani A (2023) Anomaly detection for industrial internet of things cyberattacks. Comput Syst Sci Eng 44:2361–2378. https://doi.org/10.32604/csse.2023.026712
10. Kim KH, Kim K, Kim HK (2022) STRIDE-based threat modeling and DREAD evaluation for the distributed control system in the oil refinery. ETRI J. https://doi.org/10.4218/etrij.2021-0181
11. Uncover security design flaws using the STRIDE approach. Microsoft Learn. https://learn.microsoft.com/en-us/archive/msdn-magazine/2006/november/uncover-security-design-flaws-using-the-stride-approach
12. Kitchenham B, Pearl Brereton O, Budgen D, Turner M, Bailey J, Linkman S (2009) Systematic literature reviews in software engineering—A systematic literature review
13. Fu JS, Liu Y, Chao HC, Bhargava BK, Zhang ZJ (2018) Secure data storage and searching for industrial IoT by integrating fog computing and cloud computing. IEEE Trans Ind Inform 14:4519–4528. https://doi.org/10.1109/TII.2018.2793350
14. Martins T, Oliveira SVG (2022) Enhanced modbus/TCP security protocol: authentication and authorization functions supported. Sensors 22. https://doi.org/10.3390/s22208024
15. Zada Khan W, Khan K (2019) Advanced persistent threats through industrial IoT on oil and gas industry advanced lightweight authentication protocols view project personal view project
16. Stellios I, Kotzanikolaou P, Psarakis M (2019) Advanced persistent threats and zero-day exploits in industrial internet of things. In: Advanced sciences and technologies for security applications. Springer, pp, 47–68
17. Sinhgad Institute of Technology, Panchal A, Khadse V, Mahalle P (2018) Security issues in IIoT: a comprehensive survey of attacks on IIoT and its countermeasures. In: 2018 IEEE global conference on wireless computing & networking : GCWCN-2018 : proceedings. 23–24 Nov 2018, Lonavala, India
18. Leander B, Causevic A, Hansson H (2019) Cybersecurity challenges in large industrial IoT systems. In: Proceedings, 2019 24th IEEE international conference on emerging technologies

and factory automation (ETFA) . Paraninfo Building, University of Zaragoza, Zaragoza, Spain, 10–13 Sept 2019
19. Sukiasyan A, Badikyan H, Pedrosa T, Leitao P (2022) Secure data exchange in Industrial Internet of Things. Neurocomputing 484:183–195. https://doi.org/10.1016/j.neucom.2021.07.101
20. Park S, Youm H-Y (2022) Security and privacy threats and requirements for the centralized contact tracing system in Korea. Big Data Cogn Comput 6. https://doi.org/10.3390/bdcc6040143
21. Bakhshi Z, Balador A, Mustafa J (2018) Industrial IoT security threats and concerns by considering Cisco and Microsoft IoT reference models. In: 2018 IEEE wireless communications and networking conference workshops (WCNCW), 15–18 Apr 2018
22. Mauri L, Damiani E (2022) Modeling threats to AI-ML systems using STRIDE. Sensors 22. https://doi.org/10.3390/s22176662
23. Shin DH, Kim GY, Euom IC (2022) Vulnerabilities of the open platform communication unified architecture protocol in industrial Internet of Things operation. Sensors 22. https://doi.org/10.3390/s22176575
24. AbuEmera EA, ElZouka HA, Saad AA (2022) Security framework for identifying threats in smart manufacturing systems using STRIDE approach. In: 2022 2nd International conference on consumer electronics and computer engineering (ICCECE), pp 605–612
25. Ankele R, Marksteiner S, Nahrgang K, Vallant H (2019) Requirements and recommendations for IoT/IIoT models to automate security assurance through threat modelling, security analysis and penetration testing. In: ACM international conference proceeding series. association for computing machinery
26. Borgiani V, Moratori P, Kazienko JF, Tubino ERR, Quincozes SE (2021) Toward a distributed approach for detection and mitigation of denial-of-service attacks within industrial Internet of Things. IEEE Internet Things J 8:4569–4578. https://doi.org/10.1109/JIOT.2020.3028652
27. Asif Md R al, Hasan KF, Islam MZ, Khondoker R (2022) STRIDE-based cyber security threat modeling for IoT-enabled precision agriculture systems (2022)
28. Salim MM, Rathore S, Park JH (2020) Distributed denial of service attacks and its defenses in IoT: a survey. J Supercomput 76:5320–5363. https://doi.org/10.1007/s11227-019-02945-z
29. Sengupta J, Ruj S, das Bit S (2020) A comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT
30. Li J, Lyu L, Liu X, Zhang X, Lyu X (2022) FLEAM: a federated learning empowered architecture to mitigate DDoS in industrial IoT. IEEE Trans Ind Inform 18:4059–4068. https://doi.org/10.1109/TII.2021.3088938
31. Antrobus R, Green B, Frey S, Rashid A (2019) The forgotten I in IIoT: a vulnerability scanner for industrial Internet of Things
32. Negi R, Kumar P, Ghosh S, Shukla S (2019) Vulnerability assessment and mitigation for industrial critical infrastructures with cyberphysical test bed. Taipei, Taiwan
33. Jamai I, ben Azzouz L, Azouz Saidane L, European University Cyprus, Jāmi'ah al-Lubnānīyah al-Amīrikīyah (2020) Security issues in Industry 4.0
34. González-Granadillo G, González-Zarzosa S, Diaz R (2021) Security information and event management (SIEM): analysis, trends, and usage in critical infrastructures. Sensors 21. https://doi.org/10.3390/s21144759
35. Yan Q, Huang W, Luo X, Gong Q, Yu FR (2018) A Multi-level DDoS mitigation framework for the industrial Internet of Things. IEEE Commun Mag 56:30–36. https://doi.org/10.1109/MCOM.2018.1700621
36. Sadhu PK, Yanambaka VP, Abdelgawad A (2022) Internet of Things: security and solutions survey. Sensors 22. https://doi.org/10.3390/s22197433
37. Wazid M, Bagga P, Das AK, Shetty S, Rodrigues JJPC, Park Y (2019) AKM-IoV: authenticated key management protocol in fog computing-based internet of vehicles deployment. IEEE Internet Things J 6:8804–8817. https://doi.org/10.1109/JIOT.2019.2923611
38. Ghahramani M, Javidan R, Shojafar M (2020) A secure biometric-based authentication protocol for global mobility networks in smart cities. J Supercomput 76:8729–8755. https://doi.org/10.1007/s11227-020-03160-x

39. Mourtzis D, Angelopoulos K, Zogopoulos V (2019) Mapping vulnerabilities in the industrial internet of things landscape. In: Procedia CIRP. Elsevier B.V., pp 265–270
40. Falco G, Caldera C, Shrobe H (2018) IIoT Cybersecurity risk modeling for SCADA systems. IEEE Internet Things J 5:4486–4495. https://doi.org/10.1109/JIOT.2018.2822842
41. Alruwaili FF (2021) Intrusion detection and prevention in industrial IoT: a technological survey. In: International conference on electrical, computer, communications and mechatronics engineering, ICECCME 2021. Institute of Electrical and Electronics Engineers Inc.
42. Urquhart L, McAuley D (2018) Avoiding the internet of insecure industrial things. Comput Law Secur Rev 34:450–466. https://doi.org/10.1016/j.clsr.2017.12.004
43. Gebremichael T, Ledwaba LPI, Eldefrawy MH, Hancke GP, Pereira N, Gidlund M, Akerberg J (2020) Security and privacy in the industrial Internet of Things: current standards and future challenges. IEEE Access. 8:152351–152366. https://doi.org/10.1109/ACCESS.2020.3016937
44. Alladi T, Chamola V, Zeadally S (2020) Industrial control systems: cyberattack trends and countermeasures. Comput Commun 155:1–8. https://doi.org/10.1016/j.comcom.2020.03.007

Author Index

X.-S. Yang et al. (eds.), *Proceedings of Eighth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 693,
https://doi.org/10.1007/978-981-99-3243-6