Lecture Notes in Networks and Systems 464

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

Proceedings of Seventh International Congress on Information and Communication Technology

ICICT 2022, London, Volume 3

Springer

Lecture Notes in Networks and Systems

Volume 464

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 · Simon Sherratt · Nilanjan Dey · Amit Joshi
Editors

Proceedings of Seventh International Congress on Information and Communication Technology

ICICT 2022, London, Volume 3

Springer

Editors
Xin-She Yang
Middlesex University
London, UK

Simon Sherratt
The University of Reading
Reading, UK

Nilanjan Dey
JIS University
Kolkata, India

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-19-2393-7 ISBN 978-981-19-2394-4 (eBook)
https://doi.org/10.1007/978-981-19-2394-4

Preface

The Seventh International Congress on Information and Communication Technology will be held during February 21–24, 2022, in a hybrid mode and organized by Global Knowledge Research Foundation. The associated partners were Springer and InterYIT-IFIP, Activate Learning, City of Oxford College, UK. 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 multicultural exchange of experiences and ideas.

A great pool of more than 1100 papers were received for this conference from across 95 countries among which around 300 papers were accepted and will be presented through digital platforms during the two days. Due to the overwhelming response, we had to drop many papers in the hierarchy of the quality. Total 42 technical sessions will be organized in parallel in 4 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 organized and conceptualized 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 7th Edition of this ICICT Congress 2022.

Ahmedabad, India

Amit Joshi, Ph.D.
Organising Secretary, ICICT 2022;
Director—Global Knowledge Research
Foundation

Contents

Editors and Contributors

About the Editors

Xin-She Yang obtained his D.Phil. in Applied Mathematics from the University of Oxford, and subsequently worked at the Cambridge University and the National Physical Laboratory (UK) as Senior Research Scientist. He is currently Reader in Modeling and Optimization at Middlesex University London and Adjunct Professor at Reykjavik University (Iceland). He is also elected Bye-Fellow at the Cambridge University and IEEE CIS Chair for the Task Force on Business Intelligence and Knowledge Management. He was included in the "2016 Thomson Reuters Highly Cited Researchers" list.

Simon Sherratt was born near Liverpool, England, in 1969. He is currently Professor of Biosensors in 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. He 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 Assistant Professor in the Department of Information Technology, Techno India College of Technology, India. He has authored/edited more than 75 books with Springer, Elsevier, Wiley, CRC Press and published more than 300 peer-reviewed research papers. He is Editor-in-Chief of the *International Journal of Ambient Computing and Intelligence*; Series Co-Editor of Springer Tracts in Nature-Inspired Computing (STNIC); and Series Co-Editor of Advances in Ubiquitous Sensing Applications for Healthcare, Elsevier.

Amit Joshi is currently Director of Global Knowledge Research Foundation and also Entrepreneur and Researcher who has completed his masters and research in the areas of cloud computing and cryptography in medical imaging. He has an experience of around 10 years in academic and industry in prestigious organizations. He is an active

member of ACM, IEEE, CSI, AMIE, IACSIT-Singapore, IDES, ACEEE, NPA and many other professional societies. Currently, he is 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. He has also edited more than 40 books which are published by Springer, ACM and other reputed publishers. He has also organized 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

Ali Abdelkader American University in Cairo, Cairo, Egypt

Atheel K. Abdul Zahra Department of Computer Engineering, University of Basrah, Basrah, Iraq

Adedayo O. Adio Department of Ophthalmology, University of Port Harcourt, Port Harcourt, Nigeria

Ahmad Ahmad Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Majid Ahmadi Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada

Chowdhury Farhan Ahmed University of Dhaka, Dhaka, Bangladesh

Chichebe M. Akachukwu Ahmadu Bello University, Zaria, Nigeria

Fatemah K. Al-Assfor Department of Computer Engineering, University of Basrah, Basrah, Iraq

Israa S. Al-Furati Department of Electrical Engineering, University of Basrah, Basrah, Iraq

Khamis A. Al-Karawi University of Diyala, Baqubah, Diyala, Iraq

Asmaa Abdul-Razzaq Al-qaisi Informatics Institute for Postgraduate Studies, Baghdad, Iraq

Ali Al-Swedan General Directorate of Health Affairs, Najran, Saudi Arabia

Md. Tanvir Alam University of Dhaka, Dhaka, Bangladesh

Nouf Aldossari Department of Information Systems, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia

Abdulrahman S. Alenizi Business School, University of Salford, Manchester, UK

Amal Algefes Department of Information Systems, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia

Sajjad Ali Mahmood Alkaabi University of Baghdad, Baghdad, Iraq

Héctor Allende-Cid Escuela de Ingeniería Informática, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

Abbas Almakrami General Directorate of Health Affairs, Najran, Saudi Arabia

Mohammed Almasabi General Directorate of Health Affairs, Najran, Saudi Arabia

Saad Alsaleh General Directorate of Health Affairs, Hail, Saudi Arabia

Sulaiman Alzabin General Directorate of Health Affairs, Hail, Saudi Arabia

Abdelaziz Aman General Directorate of Health Affairs, Najran, Saudi Arabia

Víctor H. Andaluz Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador;
SISAu Research Group, Universidad Tecnológica Indoamérica, Ambato, Ecuador

Aliya Nisa Anindita Bina Nusantara University, Jakarta, Indonesia

Masayoshi Aritsugi Kumamoto University, Kumamoto, Japan

Juan Arriola-Condori Universidad Privada del Norte, Lima, Perú

Moges Ayele Asale School of Psychology, Addis Ababa University, Addis Ababa, Ethiopia

Yuspani Asemki Department of Government Administration, Jusuf Kalla School Government, University of Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Abdullah Aydeger School of Computing, Southern Illinois University, Carbondale, IL, USA

Alexander Babel Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany

Sussy Bayona-Oré Dirección de Investigación, Universidad Autónoma del Perú, Lima, Peru

Elefelious Getachew Belay School of Information Technology and Engineering, Addis Ababa Institute of Technology, Addis Ababa University, Addis Ababa, Ethiopia

Narjes Benameur Research Laboratory of Biophysics and Medical Technologies, Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, Tunis, Tunisia

Vashista Bhati Department of Technical Education, Government of Rajasthan, Jodhpur, India

Dhrubajyoti Bhuyan Assam Medical College, Dibrugarh, Assam, India

Kalyan Bhuyan Department of Physics, Dibrugarh University, Dibrugarh, Assam, India

Mircea Bojan Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Kimasha Borah Centre for Computer Science and Applications, Dibrugarh University, Dibrugarh, Assam, India

Ayoub Bourbah Department of Electrical and Telecommunication, ISET Laboratory, Kenitra, Morocco

Patrick Brandtner University of Applied Sciences Upper Austria, Steyr, Austria

Crischentian Brinza "Grigore T. Popa" University of Medicine and Pharmacy, Iaşi, Romania

Joerg Bueechl Aalen University of Applied Science, Aalen, Germany

Alexandru Burlacu "Grigore T. Popa" University of Medicine and Pharmacy, Iaşi, Romania

Michael Cabanillas-Carbonell Universidad Privada del Norte, Lima, Perú

Alexis Carbajal-Torres Universidad Autónoma del Perú, Lima, Perú

Sonia Cárdenas-Delgado Departamento de Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Christian P. Carvajal Instituto de Automática, Universidad Nacional de San Juan, San Juan, Argentina

Steeven Taipicaña Cayambe Departamento de Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Salas Cesar Peruvian University of Applied Sciences (UPC), Lima, Peru; Department of Systems and Information Engineering, National University of San Marcos (UNMSM), Lima, Peru

Olena Chaikovska Kyiv National University of Culture and Arts, Kyiv, Ukraine

Dhrubajyoti Chetia Lokapriya Gopinath Bordoloi Regional Institute of Mental Health, Tezpur, Assam, India

Darren Cheung Data, Discovery and Decision Science, Allstate NI, Belfast, UK

Varin Chouvatut Department of Computer Science and Graduate School, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand

Rodriguez Ciro Department of Systems and Information Engineering, National University of San Marcos (UNMSM), Lima, Peru

António Leça Coelho Laboratório Nacional de Engenharia Civil, Lisboa, Portugal

J. Coetzer Central University of Technology, Bloemfontein, Free State, South Africa

A. Deifalla Structural Engineering and Construction Management Department, Future University in Egypt, Cairo, Egypt

Abhishek Deokar Department of Electrical and Computer Engineering, Purdue School of Engineering and Technology Indianapolis, Indianapolis, IN, USA

Jean Diatta Department of Computer Science and Mathematics, University of La Réunion, Saint Denis, France

Bhagwan Dinkar Thorat Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon, Pune, India

Lucy Doyle School of Mathematics and Physics, Queen's University Belfast, Belfast, UK

Rui P. Duarte Polytechnic of Viseu, Viseu, Portugal;
CISeD—Research Centre in Digital Services, Lisboa, Portugal

Luay Edwar University of Information and Communication Technology, Baghdad, Iraq

Mohamed El-Sharkawy Department of Electrical and Computer Engineering, Purdue School of Engineering and Technology Indianapolis, Indianapolis, IN, USA

Rangga Aditya Elias Bina Nusantara University, Jakarta, Indonesia

Marin Fotache "Alexandru Ioan Cuza" University, Iaşi, Romania

M. A. Gagarina Financial University Under the Government of the Russian Federation, Moscow, Russia;
Institute of Psychology, Russian Academy of Sciences, Moscow, Russia

Alok Singh Gahlot MBM College, Jodhpur, India

José García Escuela de Ingeniería en Construcción, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

Tejasvi Ghanshala Graphic Era Deemed to Be University, Dehradun, India

Theodosios Gkamas Laboratory of Distributed Micro-computer Systems, Department of Mathematics, University of Ioannina, Ioannina, Greece

Ramiro Gonçalves UTAD (Universidade de Trás-Os-Montes E Alto Douro), Vila Real, Portugal

Anilkumar Kothalil Gopalakrishnan Department of Computer Science, Vincent Mary School of Science and Technology, Assumption University of Thailand, Samut Prakan, Thailand

Jenny Granizo Universidad Nacional de Chimborazo, Riobamba, Ecuador

Randa Gustiawan Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Feisal Hadi Masmali School of Electrical Engineering and Computing, The University of Newcastle, Newcastle, New South Wales, Australia;
College of Business Administration, Jazan University, Jizan, Saudi Arabia

Christian Haertel Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany

Fumiko Harada Connect Dot Ltd., Tokyo, Japan

Hassan Kasim Haridi General Directorate of Health Affairs, Najran, Saudi Arabia

Ralf Härting Aalen University of Applied Science, Aalen, Germany

Vega Hugo Department of Systems and Information Engineering, National University of San Marcos (UNMSM), Lima, Peru

Sikiru T. Humble Ahmadu Bello University, Zaria, Nigeria

Suleiman U. Hussein Ahmadu Bello University, Zaria, Nigeria

Md. Sahidul Islam University of Dhaka, Dhaka, Bangladesh

Akhdiva Elfi Istiqoh Department of Government Affairs and Administration, Jusuf Kalla School Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Hiroki Ito Graduate School of Information Science and Engineering, Ritsumeikan University, Shiga, Japan

Adrian Mihai Iuoras Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Sunita A. Jahirabadkar Department of Computer Engineering, Cummins College of Engineering for Women, Pune, India

Andreas Dyrøy Jansson Department of Computer Science and Computational Engineering, UiT The Arctic University of Norway, Narvik, Norway

Marcelo Jereb Universidad Nacional del Litoral, Santa Fe, Argentina

Margarida Jerónimo Polytechnic of Viseu, Viseu, Portugal

Yuxi Jin Nagoya Institute of Technology, Nagoya, Aichi, Japan

A. Jorge Morais Universidade Aberta, Lisboa, Portugal;
LIAAD - INESC TEC, Porto, Portugal

Pankaj Chandra Kar University of Dhaka, Dhaka, Bangladesh

Vasileios Karaiskos Laboratory of Distributed Micro-computer Systems, Department of Mathematics, University of Ioannina, Ioannina, Greece

Andrey Kharitonov Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany

Svitlana Khrushch Kyiv National University of Culture and Arts, Kyiv, Ukraine

Tia Mariatul Kibtiah Bina Nusantara University, Jakarta, Indonesia

Sotirios Kontogiannis Laboratory of Distributed Micro-computer Systems, Department of Mathematics, University of Ioannina, Ioannina, Greece

Kateryna Kotsiubivska Kyiv National University of Culture and Arts, Kyiv, Ukraine

Tesnim Kraiem Laboratory of Applied Mechanics and Engineering, National Engineers School of Tunis, University of Tunis El Manar, Tunis, Tunisia

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

Danang Kurniawan Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Salam Labidi Research Laboratory of Biophysics and Medical Technologies, Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, Tunis, Tunisia

Mauricio Loachamin-Valencia Departamento de Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Cu Vinh Loc Can Tho University, Can Tho, Vietnam

Mohammad Jafar Loilatu Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Milen Loukantchevsky University of Ruse, Ruse, Bulgaria

Zhour Madini Department of Electrical and Telecommunication, ISET Laboratory, Kenitra, Morocco

Shikha Maheshwari Manipal University, Jaipur, India

Ahmed Mohammed Raoof Mahjoob University of Baghdad, Baghdad, Iraq

Temitope Mapayi Department of Computer Systems Engineering, Faculty of Information and Communication Technology, Tshwane University of Technology, Pretoria, South Africa

Adele H. Marshall School of Mathematics and Physics, Queen's University Belfast, Belfast, UK;
Faculty of Business and IT, Ontario Tech University, Oshawa, Canada

Fatma Masmoudi Department of Information Systems, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia

Zsolt Mathe SC Tehnologistic SRL, Cluj-Napoca, Romania

Carlos Maureira Escuela de Ingeniería Informática, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile;
Facultad de Ingeniería y Negocios, Universidad de las Américas, Santiago, Chile

Bouchra Meliani Department of Electrical and Telecommunication, ISET Laboratory, Kenitra, Morocco

Israel Mendonça Kumamoto University, Kumamoto, Japan

Getachew Hailemariam Mengesha School of Information Science, Addis Ababa University, Addis Ababa, Ethiopia

Shah J. Miah Newcastle Business School, The University of Newcastle, Newcastle, New South Wales, Australia

Eka Miranda Bina Nusantara University, Jakarta, Indonesia

O. L. Mokalusi Central University of Technology, Bloemfontein, Free State, South Africa

Mohamed Moustafa American University in Cairo, Cairo, Egypt

Mohammed B. Mu'azu Ahmadu Bello University, Zaria, Nigeria

Isnaini Muallidin Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Umar M. Mustapha Ahmadu Bello University, Zaria, Nigeria

Dyah Mutiarin Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Abdulrahman Nahhas Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany

Ruchi Nanda Department of CS and IT, IIS, Deemed to Be University, Jaipur, India

G. Nel Central University of Technology, Bloemfontein, Free State, South Africa

T. A. Nestik Institute of Psychology, Russian Academy of Sciences, Moscow, Russia

Joaquim Neto Universidade Aberta, Lisboa, Portugal;
Laboratório Nacional de Engenharia Civil, Lisboa, Portugal

A. N. Nevryuev Financial University Under the Government of the Russian Federation, Moscow, Russia

Neyko Neykov Angel Kanchev, University of Ruse, Ruse, Bulgaria

Joseph Ninaquispe-Florez Universidad Autónoma del Perú, Lima, Perú

Nasimul Noman School of Information and Physical Sciences, University of Newcastle, Newcastle, NSW, Australia

Aleksandar Novakovic School of Mathematics and Physics, Queen's University Belfast, Belfast, UK;
Faculty of Business and IT, Ontario Tech University, Oshawa, Canada

Achmad Nurmandi Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Ezekiel U. Okike University of Botswana, Gaborone, Botswana

Enrique Orihuela-Espinoza Universidad Privada del Norte, Lima, Perú

Zalinda Othman Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, The National University of Malaysia, Bangi, Selangor, Malaysia

Eseoghene Ovie Ahmadu Bello University, Zaria, Nigeria

Pius A. Owolawi Department of Computer Systems Engineering, Faculty of Information and Communication Technology, Tshwane University of Technology, Pretoria, South Africa

Jan Pach Aalen University of Applied Science, Aalen, Germany

Teodor Pana Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Bhasker Pant Graphic Era Deemed to Be University, Dehradun, India

Soroosh Parsai Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada

Samaneh Pazouki School of Electrical, Computer, and Biomedical Engineering, Southern Illinois University, Carbondale, IL, USA

Danny Pérez Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Aditya Permana Bina Nusantara University, Jakarta, Indonesia

Kornprom Pikulkaew Department of Computer Science and Graduate School, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand

Lucian Nicolae Pintilie Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Filipe C. Pinto Polytechnic of Viseu, Viseu, Portugal;
CISeD—Research Centre in Digital Services, Lisboa, Portugal;
Altice Labs, Aveiro, Portugal

Matthias Pohl Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany

Iolanda Valentina Popa "Grigore T. Popa" University of Medicine and Pharmacy, Iaşi, Romania

Danang Eko Prastya Master of Government Sciences, University of Muhammadiyah Yogyakarta, Yogyakarta, Indonesia;
Inspectorate, Government Regency of Magetan, Magetan, Indonesia

Vindhi Putri Pratiwi Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Denny Ardiansyah Pribadi Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Ulung Pribadi Master of Government Sciences, University of Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Luis E. Proaño Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Kusum Rajawat Shri Karni College, Jaipur, India

Hery Frédéric Rakotomalala Department of Mathematics and Computer Science, University of Antsiranana, Antsiranana, Madagascar

Awadi Rania Research Laboratory of Biophysics and Medical Technologies, Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, Tunis, Tunisia

Nicolai Romanov "Alexandru Ioan Cuza" University, Iaşi, Romania;
SC Falcon Trading SRL, Iaşi, Romania

Rastislav Róka Slovak University of Technology, Bratislava, Slovakia

Sorin Ionuț Salcu Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Nermin M. Salem Electrical Engineering Department, Future University in Egypt, Cairo, Egypt

Eric Jean Roy Sambatra Department of Electrical Engineering, Institute of Technology, Diego Suarez, Madagascar

Md. Samiullah University of Dhaka, Dhaka, Bangladesh

Bayar Shahab University of Rome Tor Vergata, Rome, Italy

Hiromitsu Shimakawa College of Information Science and Engineering, Ritsumeikan University, Shiga, Japan

Shun Shiramatsu Nagoya Institute of Technology, Nagoya, Aichi, Japan

Prabhdeep Singh Graphic Era Deemed to Be University, Dehradun, India

Vijay Singh Rathore IIS University, Jaipur, India

Omphemetse N. Small University of Botswana, Gaborone, Botswana

Daniel Staegemann Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany

Svetlana Stefanova Angel Kanchev, University of Ruse, Ruse, Bulgaria

Paulette Parra Suárez Departamento de Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Vasile Mihai Suciu Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Nur Suhailayani Suhaimi Faculty of Computer and Mathematical Sciences, University of MARA Technology, Merlimau, Melaka, Malaysia

Ahmad Syukri Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Norbert Csaba Szekely Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Petre Dorel Teodosescu Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Le Hoang Thao Can Tho University, Can Tho, Vietnam

Maryna Tolmach Kyiv National University of Culture and Arts, Kyiv, Ukraine

María Elena Tosello Universidad Nacional del Litoral, Santa Fe, Argentina

André Totohasina Department of Mathematics and Computer Science, University of Antsiranana, Antsiranana, Madagascar

Yuliia Trach Kyiv National University of Culture and Arts, Kyiv, Ukraine

Vikas Tripathi Graphic Era Deemed to Be University, Dehradun, India

Klaus Turowski Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany

José Varela-Aldas SISAu Research Group, Universidad Tecnológica Indoamérica, Ambato, Ecuador

H. J. Vermaak Central University of Technology, Bloemfontein, Free State, South Africa

Nguyen Hoang Viet Can Tho University, Can Tho, Vietnam

Tran Hoang Viet Can Tho University, Can Tho, Vietnam

Truong Xuan Viet Can Tho University, Can Tho, Vietnam

Vandana Vijay Department of CS and IT, IIS, Deemed to Be University, Jaipur, India

Rafael Villon Universidad San Martin de Porres, Lima, Peru

Ruchi Vyas MBM College, Jodhpur, India

Mohd Ridzwan Yaakub Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, The National University of Malaysia, Bangi, Selangor, Malaysia

Daiki Yamaguchi Kumamoto University, Kumamoto, Japan

Xulu Yao Manchester Metropolitan University, Manchester, UK

Moi Hoon Yap Manchester Metropolitan University, Manchester, UK

Abitassha Az Zahra Master of Government Sciences, University of Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Yanlong Zhang Manchester Metropolitan University, Manchester, UK

Younes Zouine Department of Electrical and Telecommunication, ISET Laboratory, Kenitra, Morocco

Technologies + Design for Social Integration

Marcelo Jereb and María Elena Tosello

Abstract This work synthesizes two didactic experiences carried out in the Digital Graphics Workshop, an optional course of the School of Architecture, Design, and Urbanism of the Universidad Nacional del Litoral, based in the city of Santa Fe (Argentina). The workshop proposes a transdisciplinary look that integrates knowledge and practices of three careers—architecture, visual communication design, and industrial design—to achieve a creative appropriation of the available technological instruments, in order to contribute from the design to specific problematic situations. From the professorship, we take ICT as a means to operate in the design processes, either in the ideation, development, production and/or communication stage. Technologies are not an end in themselves but a means to enhance our practices as designers operating in a community. Thus, the question is not what can we do with the available technologies? But, in what way, can we use them to help improve the living conditions of our society?

Keywords Sociotechnical networks · Design proposals · Situated knowledge

1 Introduction

1.1 A New Paradigm

We are going through a moment of change in social, cultural, and epistemological paradigms, which implies a transformation in our conception of the world and in our way of interpreting and designing the human habitat. It is the consciousness that knows that is transformed throughout history. Knowledge changes because the subjects change. In addition, in any discipline, different results will be obtained if different indicators and technological resources are used [1].

M. Jereb (✉) · M. E. Tosello
Universidad Nacional del Litoral, 3000 Santa Fe, Argentina
e-mail: marcelojereb@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_1

Today, we understand that the form of observation influences what is observed. The idea of a unique and equal reality for all is replaced by an idea of intersubjective reality, built from multiple views and interpretations. Reality is a complex and multidimensional system, i.e. a network of processes that occur simultaneously in different directions and levels of articulation.

The new paradigm does not derive only from instrumental and technological changes but also encompasses the various dimensions of man's life in society. Thomas and Buch [2] refer to the "sociotechnical" as an indissoluble dimension. Society is not determined by technology, nor is technology determined by society. Our societies are technological, just as our technologies are social. We are sociotechnical beings.

In any case, there is no doubt that digital technologies have influenced the paradigm shift and have not only caused changes in the ways of perceiving, knowing, experiencing, and communicating reality but have also caused a great impact on languages, associated with hypertextualization, telematization, and multimedia communication, which promote relational and complex thinking.

1.2 ICT and Knowledge

Access to this new context of exchange led to the appearance of unprecedented situations in which human bonds and potentials took on new dimensions. Ascott, an artist recognized for linking art, technology, and consciousness since the 1960s, proposes the concept of cyberception as a new human capacity that implies the convergence of cognitive and perceptual processes, in which connectivity to networks.

> We are acquiring new faculties, being here and potentially anywhere else at the same time, gives us new avenues of thought and perception that expand our genetic capacities. We are mediated, empowered … [3]

In addition to bringing with it a potential that has influenced various fields of knowledge, emerging technologies have generated a process of dilation of reality from the creation of a new virtual social space time, which is intertwined with physical space.

Simultaneously, the Internet has folded space, causing trans-spatial forms of interpersonal relationship, work, leisure, and knowledge. In this expanded and hybrid context, which promotes the alternation or the simultaneous experience of physical and virtual spaces, cyberculture develops [4], and to act in it, one must be connected and know how to function in a network and through digital representations.

We have moved from the information society to the network society [5]. There are inclusive sociotechnical networks formed by the integration of interdependent dimensions that in the current context of physical-virtual presence make it possible to build knowledge by carrying out responsible interaction actions and with equal opportunities to create, teach, learn, research, dialogue, design, and produce under the participatory modality of a workshop, using the communicational, transformative, and open potential of ICT [6].

2 Our Practices

From this approach that blurs—at least in part—the boundaries between technology and human sciences, between the physical and the virtual, we established pedagogical strategies to carry out significant experiences in the Digital Graphics Workshop, seeking to collaborate with problematic situations in the regional scope of our university.

For this, we decided to partner with other academic spaces of the same university and social organizations that operate in the territory and address various problems. This implied transposing more borders: between formal and non-formal education, between different branches of design, between different fields of knowledge, developing complex strategies to design, and produce artifacts, interfaces, and representations capable of articulating the links between inhabitants, their actions, and the different dimensions of the habitat in which they develop.

We implemented two experiences applying digital media to enhance creativity. In the first one—carried out in 2018—we partnered with the Civil Association La Funda, a socio-educational space dedicated to protecting and accompanying vulnerable childhoods in three peripheral neighborhoods of Santa Fe city (Fig. 1), placing emphasis on community nutrition, comprehensive sexuality education, and popular education, from the perspective of public health and human rights.

The second experience was developed in 2021 together with the Extension Project of Social Interest "Insular Narrations" [7] and the pre-professional practice of the Degree in Occupational Therapy, who have been developing their activities in the La Boca neighborhood (Fig. 2), a coastal settlement located in the valley of flooding of the Paraná river and emerged on the side of the access channel to the port of our city.

In urban situations with very different physical characteristics, with different social actors and cultural identities, and in historical contexts close in time, but

Fig. 1 With Civil Association La Funda, 2018

Fig. 2 La Boca and its natural and cultural landscape, 2021

paradoxically different (product of the derivations of the Covid-19 pandemic), the challenge was the same. What can we contribute from digital media and the disciplines of the design family to project a more sustainable and inclusive environment? How can we act from the public university to the integration of our asymmetric societies?

2.1 Methodology and Proposals

The methodology consisted of presenting the students with the problems of the analyzed sectors and the needs of their communities, as well as a set of instruments and media related to digital graphics: Web design and programming, parametric design and digital manufacturing, graphic design and multimedia, responsive artifacts (Arduino) and augmented reality.

The assignment was to create, at the intersection of problems and means, the possible design proposals that would be addressed in interdisciplinary groups, always in coordination with a general project of the workshop that provided a framework of coherence to the works (Fig. 3).

This strategy favored exchanges and collective construction, in a process of association and negotiation between subjects and institutions [8], in which the use of collaborative platforms, social networks, and shared databases was fundamental. As part of the workshop assignment, each proposal should go through the transmedia narrative in some way, either in the conception of the idea, in its development process, or in the way of communicating it to the community.

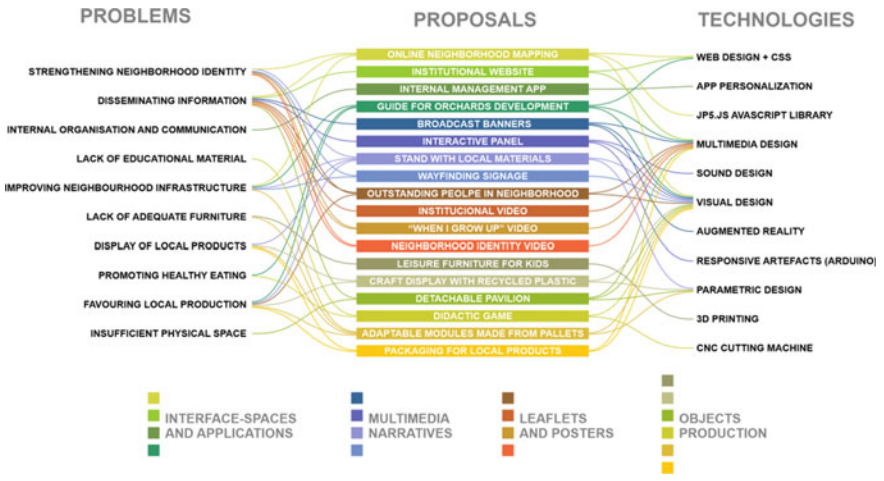


Fig. 3 Methodology, 2018–2021

Fig. 4 Proposals for La Funda: detachable pavilion, augmented reality banners, promotional videos, 2018

In the 2018 experience, productions aimed at making the association visible and contributing with its internal organization—supported with great effort by volunteer work—as well as the manufacture of objects that could collaborate in the training tasks that are developed with the children of the neighborhood. We highlight the audiovisual productions intended for advertising campaigns due to the strong affective commitment they managed to transmit, building the story from the looks and wishes of the children, some even incorporated the use of augmented reality in printed banners.

Also, interesting is the proposals that use laser cutting technology and 3D printing for the manufacture of educational games and multipurpose furniture. Particularly, striking is a proposal that combines parametric design and an artisan production/manufacturing process made with common materials, applied to the production of a removable stand (Fig. 4).

In the 2021 experience, the proposals should be aimed at making the natural and cultural heritage of Barrio La Boca visible, valuing the identity of its community, and providing innovative ideas with the capacity to drive sustainable development processes. That is why many of the proposals were related to community tourism or social economy initiatives (artisan products). In this edition, the professorship placed more emphasis on transmedia narratives, so all the proposals went through multiple platforms and languages.

We highlight some audiovisual pieces that managed to capture the particularities of the site, its history and its inhabitants to communicate them with an especially sensitive poetics. Some proposals related to local microeconomics were also very original, such as: packaging design, multifunctional furniture, and product displays, all of them raised from recycling, and the use of everyday objects from the immediate environment.

A display stand made from melted and hand-molded plastic bottle caps was highlighted. The proposal, which admits variants of use and scale, was devised based on parametric design logics, allowing the exploration of formal alternatives and combinations in order to achieve a versatile and efficient product (Fig. 5).

As a closing of the workshop, in both experiences, there were two presentation instances. An internal exhibition in which each group presented to the rest of their classmates the detected problem, the objectives of their design proposals, and the processes developed. Afterwards, there was a collaborative exposition which allowed to share the achievements with the entire educational community and, above all, to

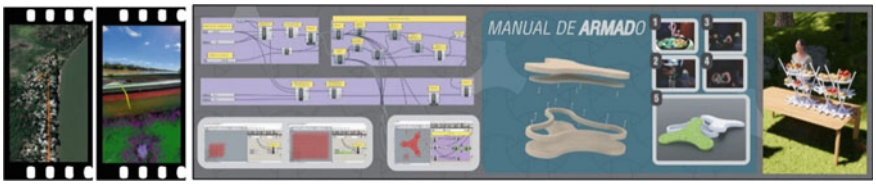


Fig. 5 Proposals for La Boca: promotional videos, display of artisan productions

Fig. 6 Final exhibition with La Funda, 2018

make what was produced in the workshop available to the associations and stakeholders (Fig. 6). Unfortunately, the restrictions derived from the Covid-19 pandemic made it impossible to carry out the final exhibition in 2021.

3 Conclusion

It is difficult to arrive at generalized assessments, given the strong link between the activities and the contextual factors of each experience. Not only is the physical and cultural environment diverse, but the members participating in the workshop—on both occasions—strongly marked the gaze that allowed the construction of lines of thought and action on which the proposals were developed.

However, we can affirm that it is highly stimulating and gratifying to see how, from the project disciplines, we can operate in the face of realities that are not infrequently ignored by the academic and professional world. We were also able to verify how design, assumed as a mediator between material means and social needs and practices, is capable of dealing with diverse problems with similar degrees of relevance.

In this way, it is not only essential to have a creative and flexible outlook, but we must learn to co-create, to think in a network, taking advantage of all available resources: subjects, objects, spaces, and media. Producing design and building knowledge from this multiple perspective, taking advantage of the communicational, transformative, and open potential of ICT, are undoubtedly the greatest achievement of these experiences.

Acknowledgements To the Universidad Nacional del Litoral and to all the people who offered their time and knowledge in these experiences.

References

1. Díaz E (2009) Seminario La Edad de la Crítica. Acerca de la ciencia y la epistemología en la vida social. Facultad de Humanidades y Artes, Rosario
2. Thomas H, Buch A (2008) Actos, Actores y Artefactos. Buenos Aires, Universidad Nacional de Quilmes
3. Ascott R (1995) The architecture of cyberception. In: Architectural design, pp 38–41
4. Lévy P (2007) Cibercultura. La cultura de la sociedad digital. Anthropos, Barcelona
5. Castells M (2010) Comunicación y poder. Alianza, Madrid
6. San Martín P (2008) Hacia un dispositivo hipermedial dinámico. Colección Música y Ciencia, N°III. Universidad Nacional de Quilmes, Buenos Aires
7. Mines P (2019) Narraciones insulares: itinerarios para un turismo sustentable en la MRI. Universidad Nacional del Litoral, Santa Fe
8. Scolari C (2008) Hipermediaciones. Elementos para una Teoría de la Comunicación. Gedisa, Barcelona

Weight-Based Dynamic Hybrid Recommendation System for Web Application Content

Margarida Jerónimo, Filipe C. Pinto, and Rui P. Duarte

Abstract This paper presents a prototype for a Web application recommendation system's content applied to movies' recommendations. It learns the pattern of user content consumption, predicting what he will consume in future based on similar items to those he has shown interest. It considers similarity with neighbor users, thus creating a user model. Content-based filtering, collaborative filtering, and memory-based on hybrid filtering techniques are used. Content-based filtering allows to extract the fundamental features or attributes of the items and select similar items. Moreover, it proposes predicted classifications for the items of interest not yet classified by the active user. Collaborative filtering allows applying the KNN methodology to identify the similarity between the active user located in the neighborhood and propose predicted classifications for items of interest not yet classified. Hybrid filtering combines the two methodologies to overcome their drawbacks. A weighted approach is applied, allowing a dynamic linear combination of collaborative and content-based filtering. The results obtained were empirically relevant in the experimental evaluation, matching with the results presented in similar studies validated with RMSE metrics.

Keywords Information systems · Recommender systems · Human centered computing · User models · Web-based interaction

M. Jerónimo · F. C. Pinto · R. P. Duarte (✉)
Polytechnic of Viseu, Viseu, Portugal
e-mail: pduarte@estgv.ipv.pt

M. Jerónimo
e-mail: guijeronimo@msn.com

F. C. Pinto
e-mail: fcpinto@estgv.ipv.pt

F. C. Pinto · R. P. Duarte
CISeD—Research Centre in Digital Services, Lisboa, Portugal

F. C. Pinto
Altice Labs, Aveiro, Portugal

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_2

1 Introduction

This paper covers the techniques of content-based filtering (CBF), collaborative filtering (CF), and hybrid filtering (HF) based on memory. These approaches have drawbacks and advantages. CBF tends to specialize the recommendations around the characteristics of the items or the active user since it does not consider the tastes of the other users. CF has the disadvantage of cold start and does not take into account the characteristics of the items. HF combines the two methodologies to overcome their drawbacks and can also follow several approaches.

This paper's contribution is a prototype of a dynamic hybrid-based recommendation system, with a Web interface, for movie recommendations. This makes it possible to experiment, test, and evaluate recommendation algorithms based on CBF, CF (memory-based), and HS, which allows an identified user to obtain a good recommendation for movies. The system uses offline and online data sources to proceed with the prediction/recommendation. The system is implemented using Django, a python-based Web platform for the development, and deployment of Web applications. It has access to online movie data sources to present the user with information and sources for CBF made available by the TMDb Website [1]. This system implements a database of evaluations of identified users, combined with online and offline sources, to improve the recommendations [2].

This paper is organized as follows. In Sect. 2, an overview of the most relevant techniques available in the literature is presented. Section 3 explains the steps taken to develop the hybrid system. Section 4 analyzes and discusses the results obtained. Finally, Sect. 5 presents the conclusions and points out future work.

2 Background

RS has been gaining prominence in several areas of research in the last decades. However, it is possible to improve and create new algorithms, more efficient and effective in recommendations. According to Resnick et al. [3], RS increases the capacity and effectiveness in decision-making, and the improvement of the use of efficient and accurate recommendation techniques is very important for a system, which provide useful recommendations for its users [4].

RS is classified based on the application domains used, how they formulate recommendations and the algorithms they implement [5]. RS can be classified based on content [6], collaborative filtering [7], and hybrid [8]. Several authors complement the classification of RS in demographic data [8], knowledge [9], and community [10], in order to refine its characterization. CBF uses prediction and training techniques collected from users and recommends items based on the textual information of an item, used to construct item profiles [11]. User profiles can be defined by constructing a model of user preferences, using items, in which the user is interested [12]. CF is used for systems in which information filtering is carried out with the intervention

of users. This system focuses on the interaction between user and item, ignoring the characteristics of the items and worrying about which items have already been consumed by the user. CF techniques create a matrix of users with their preferences items. They combine users with relevant preferences, calculating similarities between their profiles to make recommendations by creating a neighborhood. Users receive recommendations for items classified by others in their neighborhood [13]. CF has serious limitations in the quality of recommendations, such as the sparsity problem, cold start problem, scalability, fraud, black sheep, gray sheep, and the synonymy [14].

HF combines different techniques, which aim to bridge the limitations of each method. The hybrid filtering system can be implemented in several ways, such as making predictions based on content filtering and based on CF, separately and combining predictions, adding content-based features to a collaborative approach (and vice versa) or unifying approaches in a model [4]. For RS to be effective, they need to learn from user interaction with the system. A new user or a new item is a cold start problem. For a new user, it is recommended that the system first extracts his preferences. Next, the system selects the group of similar users (neighbors) who rated items and thus, makes the recommendation [15, 16].

3 Development of a RS for Movies

This section features the RS developed for movie recommendations. It makes it possible to experiment, test, and evaluate recommendation algorithms based on CBF, CF based on memory, and on dynamic HF. Moreover, the RS allows an identified user to obtain a recommendation for movies not yet evaluated.

3.1 Prototype

RS features an interface that allows any user to obtain statistics, graphics, movie information, and non-personalized Top-N recommendations of the most popular movies from MovieLens offline datasets, provided by GroupLens [17]. We implement the CF algorithm that applies IMDb's equation to define the "TopRated 250titles". Authenticated users (after registration in the system) can get customized recommendations for collaborative, content-based, or hybrid filtering. Although in the MovieLens dataset, the user is characterized by his userId, other characteristics are safeguarded: name, email address, and demographic data.

The Web interface is developed in Django and allows an unauthenticated user to obtain information about movies, create a user, or authenticate to classify movies. When the user authenticates, the main page will display a recommendation of type Top-N, applying the dynamic weighted hybrid technique [18]. An unidentified user can use the system to search MovileLens datasets offline for information on movies,

statistics, and graphs of prepared and clean data. The user can also obtain a recommendation based solely on the textual similarity of movies, or get a simple, non-personalized collaborative recommendation from top popularity of movies. On the other hand, an authenticated user can obtain personalized recommendations based on the offline MovieLens dataset, enhanced with data from database ratings. For the preparation and cleaning of the offline MovieLens dataset, graphic elaboration and support of recommendation algorithms, the libraries Python pandas and NumPy, were applied.

3.2 Types of Recommendations

Non-personalized Recommendations are of the top-N type, with possible classification prediction. Their filtering does not require the identification of the user. Before obtaining information about movie titles that approximate the search, the fuzzy string matching [19] technique is used. Non-authenticated users can also search for similar movies based on the content (genre or title). For this purpose, the cosine similarity is used, applied to the term frequency-inverse document frequency (TF-IDF) [20]. Term frequency (TF) is the frequency of a word (or term, t) in the current document (d), in relation to the total number of words of that document (N_d) and is given by $TF_{t,d} = \sum_{i=0}^{N_d} t_i / N_d$. To assign higher ratings to rare terms in documents, it is used the inverse document frequency. It is given by $\text{IDF}_t = \log_{10}(N / \sum_{i=0}^{N} t_i)$ where N is the number of documents. To obtain the top popular movies, a CF technique was implemented that allows to solve the cold start problem of new users in the system. It considers a weighted classification for each item (classification prediction based on popularity) and is given by $\text{WR} = v/(v+m) \times R + m/(v+m) \times C$ where v is the number of votes in the movie, m refers to the minimum number of votes to be considered in the "top rated" list, R is the average rating of the movie, and C is the average of all average ratings. The choice of the value of m implies removing items that have a number of ratings lower that a threshold for m. By calculating the popularity of all items, a top-N recommendation is obtained. This method is not sensitive to the interests and tastes of specific users.

Personalized Recommendations relate to recommendations of type top-N with classification forecast, whose filtering requires the identification of the user, and results obtained generate a top-N list. The prediction engines use other users' (or item) ratings or content information. It then predicts how the user would like a particular item. Another way to make recommendations is to deliver an ordered list of items [2].

The learning of each user's interests is conducted based on the set of evaluations carried out and can thus assume with some degree of confidence that the user shows interest in items with identical characteristics. In this work, the genres of movies were considered. TF-IDF technique is also implemented to determine the similarity between users and items. By determining the internal product between the by movies'

profiles and the vector of ratings, the TF-IDF vector is obtained, which constitutes the user's profile after normalization. The cosine similarity is used and is given by

$$sim(U, I) = \frac{\sum_{j=1}^{t} \left(\text{TF-IDF}_{j,U} \times \text{TF-IDF}_{j,I} \right)}{\sqrt{\sum_{j=1}^{t} \left(\text{TF-IDF}_{j,U} \right)^2} \times \sqrt{\sum_{j=1}^{t} \left(\text{TF-IDF}_{j,I} \right)^2}} \tag{1}$$

The resulting similarities should range from 0 to 1. If that occurs, $(U, I) = 0$, which implies that the two profiles are independent; otherwise, the two profiles have similarities. To obtain a classification prediction, to an item that has not yet been classified (given by I_a), the following relation applies: $Pred(U_i, I_a) = \sum(sim(U_i, I_b) \times r_{U_i, I_b}) / \sum sim(U_i, I_b)$, where $sim(U_i, I_b)$ uses similarities to the items already sorted and the ratings assigned to those items (r_{U_i, I_b}).

In the case of collaborative recommendation, top-N movies are calculated based on the similarity between users for an active user. The process is divided into the phases of calculating similarities between users, determining neighborhoods for users, and predicting ratings for items not yet classified by the active user [21]. The predicted classification is given by

$$P_{u,i} = \overline{r_u} + \frac{\sum_{v \in V} sim(r_u, r_v) \times (r_{v,i} - \overline{r_v})}{\sum_{v \in V} |sim(r_u, r_v)|} \tag{2}$$

where i corresponds to the item for which a prediction is to be determined for a user, u, that has not yet classified an item; $r_{v,i}$ is the classification given by the user v to the item i, and V represents the neighborhood of user u.

Several techniques based on content and collaborative approaches are used to obtain movies to recommend. For each method, weights are assigned, and a weighted HF system is obtained. It allows getting the top-N preferred movies of an authenticated user, including the predictions of classifications. Following a dynamic weighted approach [2], it is given by $P_{u,i} = \alpha_1 \times P^1_{u,i} + \cdots + \alpha_n \times P^n_{u,i}$. Here, α is adjusted according to each technique used. Since most hybrid RS follow CF, CBF, or a combination of both, these techniques will be considered exclusively. Thus, the confidence α_u of the prediction of CF ($P^{CF}_{u,i}$) should increase when the number of items evaluated by user u increases. The confidence of the system based on content $P^{CBF}_{u,i}$ is given by $1 - \alpha_u$. In this paper, both recommendation techniques are combined, the value of the predicted score is given by $P_{u,i} = \alpha_u \times P^{CF}_{u,i} + (1 - \alpha_u) \times P^{CBF}_{u,i}$. According to Cremonesi et al. [2], α_u is given by $\alpha_u = \frac{t_u}{k} \times 0.9$, where $k \in \mathbb{N}$ is the number of used neighbors in the collaborative classification and t_u is the number of items of $I \setminus I_u$.

The evaluation metrics used to validate the algorithms implemented were statistical precision methods that assess the accuracy of a filtering technique by comparing the predicted classifications directly with the actual user classification. The root mean square error (RMSE) is used and measures the accuracy of the prediction of

the ratings [5]. For the construction of the system, the approaches described in this section are implemented individually and combined, which allows the construction of a hybrid weighted system.

4 Analysis of Personalized Movie Recommendation

For the tests and analysis of the results of the personalized recommendations, the ratings dataset was subdivided into training and testing. This allows to make recommendations on the training set and compare the algorithms with the test set. The variable x is used to determine the percentage of data used in the training and test sets. A value of $x = 0.8$ indicates that 80% of the data is used for training and 20% for tests.

4.1 Recommendation for Items Profiles and Users

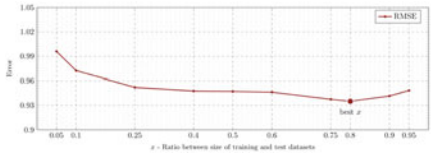
The profiles of users are created based on the ratings they made and are a part of the training dataset. By introducing variations in the x value, Fig. 1a shows that the most convenient value for x is 0.8. It is the lowest value of x, which indicates better precision.

According to Do et al. [18], the CBF algorithm presents an RMSE ≈ 1.2, for $x = 0.8$. Our results present an RMSE ≈ 0.938, for $x = 0.8$.

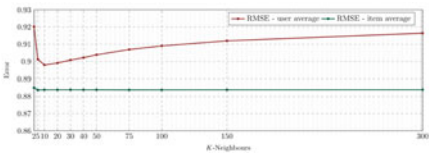
4.2 Collaborative Recommendation Based on Users Similarity

The sensitivity of prediction errors was tested with $x = 0.8$, in relation to the size of the neighborhood. Results are presented in Fig. 1b. Regarding data rearrangement due to lack of ratings in the user-item matrix, the average of the user's adjusted rating (user average in Fig. 1b), the number of neighbors to consider for the recommendations varies between 10 and 40.

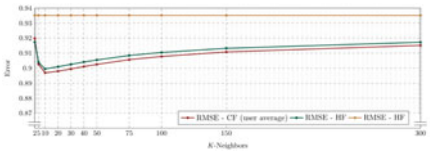
As depicted in Fig. 1b, the system is not sensitive to the number of neighbors. Moreover, the implementation of the item average produced better accuracy of the CF recommendation system. For 10 to 40 neighbors, RMSE ≈ 0.9 for the user's average, and RMSE ≈ 0.884 to the item's average. Additionally, considering the average of adjusted ratings of the item to fill the ratings not yet made, the recommendations are better and not sensitive to the number of neighbors.



(a) Sensitivity to the parameter x.



(b) Sensitivity to the size of neighbors.



(c) Comparison of RMSE amongst filtering techniques versus number of neighbors.

Fig. 1 RMSE sensitivity

4.3 Hybrid Recommendation

The implemented technique always considered a confidence factor of $\alpha_u = 0.9$ for CF predictions. All users have rated at least 14 items. Thus, with the exception of tests carried out with 2 neighbors, the results show that the HF implementation follows the CF implementation, with a slightly higher error. Therefore, the CBF implementation, does not contributes to the improvement of recommendations, as shown in Fig. 1c. However, when considering only 2 neighbors, the CBF implementation contributes to the improvement of recommendations.

Results show that CBF recommendation systems are good at finding similar items and do not provide relevant information on the ratings of the active user. They are relevant in scenarios where the user does not have many preferences in his profile. Therefore, the cold start problem is a reality that can be overcome with the contribution of CBF techniques. However, in the used dataset and with the HF implementation, this evidence only is valid for 5 or less neighbors. Considering the results presented in Do et al. [18], RMSE errors always exceed 1.1, while in this study, all were bellow 0.94.

5 Conclusions

This paper shows that the MovieLens datasets allowed to validate the trends already presented in similar studies. However, due to their size, some conditions of the application of the techniques were not completely clarified. Specifically, based on the datasets used, since users have always rated at least 20 movies, there was no significant contribution from CBF to the hybrid recommendation, except when considering

two neighbors. Moreover, the tests carried out on the prototype and validation results, the obtained recommendations provide values for RMSE below 0.94 in all implemented techniques.

Future work will follow several directions. First, the MovieLens datasets will be combined with other similar datasets (e.g., Netflix, TMDb, and IMDb). Since this paper only applies memory-based techniques, it is essential to use model-based methods such as machine learning. Another important work to be carried out is related to the long tail effect and apply precision metrics (like serendipity[22]) to support decision.

Acknowledgements This work is funded by National Funds through the FCT—Foundation for Science and Technology, I.P., within the scope of the project Ref. UIDB/05583/2020. Furthermore, we would like to thank the Research Center in Digital Services (CISeD) and the Polytechnic of Viseu for their support.

References

1. TMDb: The movie database (tmdb) (2021). URL https://www.themoviedb.org/
2. Cremonesi P, Koren Y, Turrin R (2010) Performance of recommender algorithms on top-n recommendation tasks. In: Proceedings of the fourth ACM conference on Recommender Systems, RecSys '10. Association for Computing Machinery, New York, NY, USA, pp 39–46. https://doi.org/10.1145/1864708.1864721
3. Resnick P, Varian HR (1997) Recommender systems. Commun ACM 40(3):56–58
4. Burke R (2002) Hybrid recommender systems: survey and experiments. User Modeling User-Adapted Interaction 12(4):331–370. https://doi.org/10.1023/A:1021240730564
5. Shani G, Gunawardana A (2011) Evaluating recommendation systems. In: Recommender systems handbook. Springer, Heidelberg, pp 257–297 https://doi.org/10.1007/978-0-387-85820-3_8
6. Lang K (1995) Newsweeder: Learning to filter netnews. In: Prieditis A, Russell S (eds) Machine learning proceedings 1995. Morgan Kaufmann, San Francisco (CA), pp 331–339. https://doi.org/10.1016/B978-1-55860-377-6.50048-7
7. Resnick P, Iacovou N, Suchak M, Bergstrom P, Riedl J (1994) Grouplens: an open architecture for collaborative filtering of netnews. In: Proceedings of the 1994 ACM conference on Computer Supported Cooperative Work, CSCW '94. Association for Computing Machinery, New York, NY, USA, pp 175–186. https://doi.org/10.1145/192844.192905
8. Balabanović M, Shoham Y (1997) Fab: content-based, collaborative recommendation. Commun ACM 40(3):66–72. https://doi.org/10.1145/245108.245124
9. Gennari JH, Musen MA, Fergerson RW, Grosso WE, Crubézy M, Eriksson H, Noy NF, Tu SW (2003) The evolution of protégé: an environment for knowledge-based systems development. Int J Human-Comput Stud 58(1):89–123. https://doi.org/10.1016/S1071-5819(02)00127-1
10. Pham MC, Cao Y, Klamma R, Jarke M (2011) A clustering approach for collaborative filtering recommendation using social network analysis. J UCS 17(4):583–604
11. Herlocker JL (2000) Understanding and improving automated collaborative filtering systems. Citeseer
12. Ghazanfar M, Prugel-Bennett A (2010) An improved switching hybrid recommender system using naive bayes classifier and collaborative filtering. In: The 2010 IAENG international conference on data mining and applications (19/03/10). https://eprints.soton.ac.uk/268483/. Event Dates: 17–19 Mar 2010

13. Isinkaye F, Folajimi Y, Ojokoh B (2015) Recommendation systems: principles, methods and evaluation. Egyptian Informatics J 16(3):261–273. https://doi.org/10.1016/j.eij.2015.06.005
14. Kim HN, Ji AT, Ha I, Jo GS (2010) Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. Electronic Commerce Res Appl 9(1):73–83. https://doi.org/10.1016/j.elerap.2009.08.004. Special Issue: Social Networks and Web 2.0
15. Sarwar B, Karypis G, Konstan J, Riedl J (2001) Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th international conference on world wide web, WWW '01. Association for Computing Machinery, New York, NY, USA, pp 285–295. https://doi.org/10.1145/371920.372071
16. Huang Z, Chen H, Zeng D (2004) Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. ACM Trans Inf Syst 22(1):116–142. https://doi.org/10.1145/963770.963775
17. GroupLens: Movielens database (2018). https://grouplens.org/datasets/movielens/
18. Do H, Le T, Yoon B (2020) Dynamic weighted hybrid recommender systems. In: 2020 22nd International Conference on Advanced Communication Technology (ICACT), pp 644–650. https://doi.org/10.23919/ICACT48636.2020.9061465
19. Gonzalez J (2020) Fuzzy wuzzy. https://github.com/seatgeek/fuzzywuzzy
20. Christian H, Agus MP, Suhartono D (2016) Single document automatic text summarization using term frequency-inverse document frequency (tf-idf). ComTech: Comput Math Eng Appl 7(4):285–294. https://doi.org/10.21512/comtech.v7i4.3746
21. Ekstrand MD, Riedl JT, Konstan JA (2011) Collaborative filtering recommender systems. Found Trends Hum-Comput Interact 4(2):81–173. https://doi.org/10.1561/1100000009
22. McNee SM, Riedl J, Konstan JA (2006) Being accurate is not enough: how accuracy metrics have hurt recommender systems. In: CHI '06 extended abstracts on human factors in computing systems, CHI EA '06. Association for Computing Machinery, New York, NY, USA, pp 1097–1101. https://doi.org/10.1145/1125451.1125659

Measuring the Success of the Ngawi District Government Web site Using the Delone and Mclean Model

Danang Eko Prastya, Ulung Pribadi, and Abitassha Az Zahra

Abstract The purpose of this study is to determine the Web site's success in providing public information for the Ngawi Regency Local Government. The Delone and Mclean approach is used to evaluate a Web site as an information system. The study employs a quantitative methodology. The data were gathered by the distribution of questionnaires to 100 respondents. Structural equation modeling (SEM) and Smart PLS 3.0 partial least squares (PLSs) software were used to analyze the data. The findings indicated that all indicators were deemed genuine, and the variables had a high degree of reliability—information quality has an effect on Web site-user happiness. Meanwhile, service quality has a favorable influence on Web site users' use and satisfaction. Meanwhile, the quality of the system has minimal influence on its usability or user satisfaction.

Keywords Web site information quality · Public service · E-government service

1 Introduction

The government's responsibility is to provide services to the people with accurate and concise information [1]. The development of e-government takes place on four levels, one of which is the creation of a public-accessible Web site [2]. The availability of the Web site is regarded as clear proof of the government's sincerity in adopting e-government. As a result, creating an information Web site for government institutions is a key first step toward e-government adoption. A Web site becomes crucial in providing information to the public as the first stage in establishing e-government. A Web site can inform the public about all government operations and policies. Additionally, increasing community interest in and make it easy accessibility to public

D. E. Prastya (✉) · U. Pribadi · A. A. Zahra
Master of Government Sciences, University of Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: danang.eko.psc20@mail.umy.ac.id

Present Address:
D. E. Prastya
Inspectorate, Government Regency of Magetan, Magetan, Indonesia

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_3

services. The Web site can serve as a repository for the government's wealth of information pertaining to the development and implementation of electronic government [3]. It promotes not just public interest in using the Web site to communicate with the government but also user comfort [4]. According to a recent survey, the average page speed rating for 30 Indonesian local government Web sites is just 59%. This indicates that the Web sites of local governments are of low quality. The quality of local government Web sites continues to be deficient, according to Google's index. The Web site's inadequate quality will very certainly deter Internet users from visiting government Web sites [5].

Meanwhile, transparency and accountability on government Web sites are critical in giving information about government performance. It is possible to create a Web site by standardizing the contents of the minimal Web site needs. Hence, this study wants to measure the effectiveness of an Ngawi government Web site in increasing public interest.

2 Use of Web site in Government

E-government is the government's use of technology to streamline the management of information and services [6]. E-government evolved as a result of developments in information and communication technology (ICT), with the goal of offering government services to stakeholders [7]. Three pillars underpin e-government: information technology, stakeholders, and outcomes. The outcomes component is based on the electronic services provided to (1) strengthen and transform stakeholder interactions and relationships, (2) empower citizens through transparent information access, and (3) improve the efficiency and effectiveness of government activities by reducing corruption and costs while increasing transparency, accessibility, and earnings [8].

Often, e-government services are supplied through government-created and administered Web sites [9]. Government online portals provide a variety of services, with some focused only on information access [10]. According to the Delone and Mclean model, the components of system quality, information quality, and service quality all contribute to user satisfaction and ultimately result in net benefits or end outcomes [11, 12]. The quality of the system may be defined as its performance, which relates to the hardware, software, regulations, and processes of the information system's capacity to meet the demands of users [13]. Usability, accessibility speed, system dependability, flexibility, and the utility of features and functions are all considered to be system quality aspects [11]. The quality of information produced by a system is a desirable attribute [14]. Users will be happier with high-quality information, but low-quality information will need consumers to take more effort to receive the information they desire [15]. The term "service quality" refers to the level of assistance provided by a system's development to its users [11]. When services meet users' expectations, user satisfaction increases. Satisfaction is created when the provider satisfies the customer's wishes [16]. The model's components [11] can be used to assess the success of the government's e-government implementation.

A high-quality government Web site can help build trust and ensure the successful deployment and delivery of e-government services [11, 12]. Additionally, several studies demonstrate that the Delone and McLean model may be utilized to evaluate the success of electronic governance in municipal government [15].

3 Research Method

This research, which applies quantitative approaches, examines the official Web site of the Ngawi district government. Data collection occurred via the distribution of questionnaires to study participants. The research respondents were drawn from a sample of the 870,057-strong Ngawi community. The Slovin formula was used to calculate a sample size of 100 respondents with a margin of error of 10%. The obtained data were then analyzed using the Smart PLS 3.0 partial least square (PLS) software, which is one of the approaches in variance-based structural equation modeling (SEM) statistics that was developed to handle multiple regression when specific data issues such as small sample sizes, missing data, and multicollinearity occurred [15]. The Delone and Mclean models are used to determine the performance of the Web site for the Ngawi Regency Government. The researcher then produced final results or benefits by using the Delone and Mclean model's independent variables of information quality, system quality, and service quality, as well as the dependent variables of use intensity and user satisfaction. The following provides an overview of the conceptual framework.

Based on the framework of thinking (Fig. 1), the following hypothesis can be put forward:

H1a. The quality of information has a strong favorable effect on its utilization.

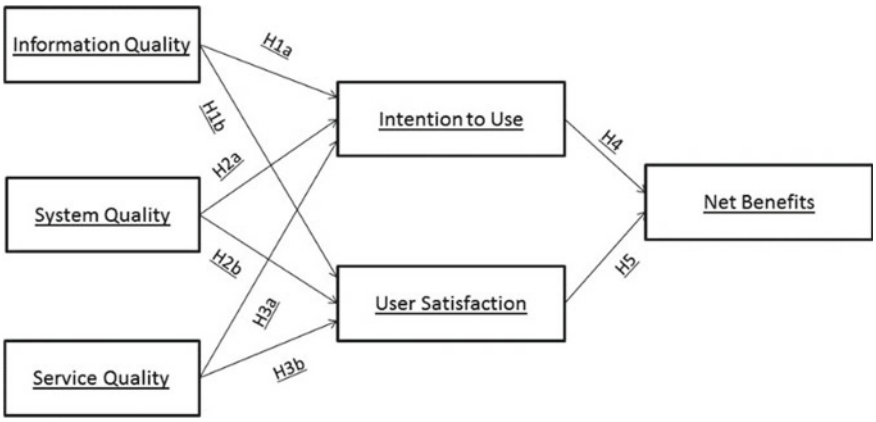


Fig. 1 Conceptual framework

H1b. The quality of information has a measurable beneficial effect on customer happiness.
H2a. The quality of the system has a favorable and substantial effect on its use.
H2b. The quality of the system has a positive and substantial impact on user satisfaction.
H3a. The quality of service has a significant and beneficial effect on system use.
H3b. The quality of service has a significant and beneficial effect on consumer happiness.
H4. The degree of utilization has a significant positive effect on the net outcomes gained.
H5. User experience has a significant beneficial effect on the final product.

4 Results

4.1 Respondent Profile

The number of persons who participated in this study was 100, as indicated in Table 1.

According to Table 1, there were 56 males and 44 women among the 100 responders, with 55% of those aged 20–30 years old. Meanwhile, at the education level, undergraduate level 1 education accounted for 56%, while SMA/SMK/MA education accounted for 41%. As a result, the respondents in this study are diverse in terms of gender, age, and education.

Table 1 Research respondent profile

Category	Amount	Percentage
Gender		
Male	56	56
Female	44	44
Age (years)		
20–30	55	55
31–40	27	27
41–50	18	18
Education		
Senior high school	41	41
Undergraduate	56	56
Graduate	3	3

4.2 Outer Model Test Results

The reflective outer model is often used to illustrate how the manifest variable or observed variable reflects the latent construct being evaluated, especially by doing confirmatory factor analysis on the indicators composing the latent construct. By examining the value of the outer loading, a convergent validity test is done to determine the indicators' validity. Table 2 shows the outer loading value.

The value of the rule of thumb, as shown in Table 2, is the data validity test for convergent decision requirements. The value of the outer loading is acceptable at 0.50, but the value of the outer loading is preferable at 0.7. If the resulting outer loading value is less than 0.50, the indicator is deemed invalid and is deleted from the research instrument. As shown in Table 2, all indicators have an outside loading value of 0.50, suggesting that they are genuine. Additionally, the extracted average variance (AVE) value indicates convergent validity. At a minimum, the AVE value must be 0.50. The AVE value is as follows.

As seen in Table 3, all variables have a value larger than 0.50. All variables with good convergent validity account for more than half of the variance in each indicator on average. The dependability of the variables in this research was determined using

Table 2 Value of outer loadings indicator

Indicator	Intensity of use	Benefits	User satisfaction	Information quality	Service quality	System quality
IQ1				0.863		
IQ2				0.910		
IQ3				0.863		
IU1	0.920					
IU2	0.925					
IU3	0.916					
NB1		0.890				
NB2		0.920				
NB3		0.858				
SQ1						0.904
SQ2						0.930
SQ3						0.871
ServQ1					0.863	
ServQ2					0.914	
ServQ3					0.862	
US1			0.857			
US2			0.910			
US3			0.880			

Source Analyze by Author, 2021

Table 3 Average variance extracted (AVE) value

Variable	Average variance extracted (AVE)
Intensity of use	0.847
Benefits	0.792
User satisfaction	0.779
Information quality	0.773
Service quality	0.774
System quality	0.813

Table 4 Cronbach's alpha value

Variable	Cronbachs alpha
Intensity of use	0.910
Benefits	0.868
User satisfaction	0.856
Information quality	0.853
Service quality	0.854
System quality	0.885

Cronbach's alpha and composite dependability ratings. Cronbach's alpha is a statistic that indicates the maximum value of a construct. Cronbach's alpha higher than 0.60 suggests the trustworthiness of a construct. Cronbach's alpha is calculated as follows in this inquiry.

Meanwhile, composite reliability is used to identify the true value of a variable. As shown in Table 4, all variables have a Cronbach's alpha value larger than 0.60, suggesting that they meet reliability criteria. The data are considered reliable if the composite reliability rating is greater than 0.70. The following is a summary of the composite reliability processing results. As shown in Table 5, all variables have a composite dependability score of greater than 0.70, indicating that they meet the reliability criteria.

Table 5 Composite reliability value

Variable	Composite reliability
Intensity of use	0.943
Benefits	0.919
User satisfaction	0.913
Information quality	0.911
Service quality	0.911
System quality	0.929

Table 6 R-square value

Variable	R Square
Intensity of Use	0.546
Benefits	0.745
User satisfaction	0.613

4.3 Result of Structural Model Testing (Inner Model)

The inner model assessment aims to determine if factors in the study have a direct or indirect effect. Additionally, this is done to determine whether the model being used is capable of resolving existing problems. The data from the R-square and path coefficient are utilized to validate the structural model (hypothesis testing). The R-square value indicates the extent to which variables in the model have an effect on one another. For instance, Table 6 gives the R-square values for the following.

According to Table 6, the mediating variable, utilization intensity, has an R-square of 0.546, or 54%. This suggests that the degree to which the variables information quality, system quality, and service quality impact the intensity of use accounts for 54% of the variance, with the remaining 46% explained by other factors. Meanwhile, another moderating variable, customer satisfaction, is 0.613, or 61%. This condition illustrates that the degree to which information, system, and service quality affect consumer satisfaction accounts for 61% of customer happiness. Other variables, on the other hand, have an effect on the remaining 39%. On the other hand, the user variable R-square has a value of 0.745, or 75%. As a result, the variable intensity of usage and user satisfaction accounts for 0% of the variance.

Smart PLS tests hypotheses using the bootstrapping method. Bootstrapping is a resampling test that a computer system does to ascertain the accuracy of a sample estimate. It is used to determine whether or not there is a statistically significant relationship between the observed variables. Utilize the amount of the hypothesis support to compare the table and statistic data. If statistics can be believed. The hypothesis is supported if the value of table is larger than zero. If the level of confidence is 95% (alpha is 5%), the table has a value of 1.96. The hypothesis is said to be accepted if the statistic is more than 1.96; it is said to be rejected if the value is less than 1.96 [6]. Table 7 and Fig. 2 summarize the results of the hypothesis testing.

Table 7 and Fig. 2 show that there are four approved hypotheses and four rejected hypotheses for hypothesis testing. H1b (t-statistical value 4.126 > 1.96), H3a (t-statistical value 4.725 > 1.96), H3b (t-statistical value 2.345 > 1.96), and H5 (t-statistical value 8.906 > 1.96) are the acceptable hypotheses, while the rejected hypotheses are H1a (t-statistical value 1.656.96), H2a (t-statistical value 0.416.96), H2b (t-statistical value 0.605.96), and H4 (t-statistical value 0.605.96). The following is the accepted hypothesis:

H1b. The quality of information has a significant positive effect on customer happiness.

Table 7 Path coefficients bootstrapping

	Original sample (0)	Sample mean (M)	Standard error (STERR)	T statistics (|0/STERR)	P values
Intensitas Penggunaan → Kemanfaatan	0.148	0.139	0.087	1.708	0.091
Kepuasan Pengguna → Kemanfaatan	0.744	0.750	0.084	8.906	0.000
Kualitas Informasi → Intensitas Penggunaan	0.193	0.188	0.116	1.656	0.101
Kualitas Informasi → Kepuasan Pengguna	0.484	0.487	0.117	4.126	0.000
Kualitas Layanan → Intensitas Penggunaan	0.527	0.530	0.111	4.725	0.000
Kualitas Layanan → Kepuasan Pengguna	0.278	0.256	0.119	2.345	0.021
Kualitas Sistem → Intensitas Penggunaan	0.066	0.061	0.160	0.416	0.678
Kualitas Sistem → Kepuasan Pengguna	0.080	0.098	0.131	0.605	0.547

H3a. The quality of service will have a positive and substantial influence on how the system is used.
H3b. Service quality will have a significant positive impact on user satisfaction.
H5. Satisfaction with the user has a substantial beneficial influence on the net outcomes received.

While the rejected hypotheses are as follows:

H1a. The quality of information will have a significant positive effect on.
H2a use. The quality of the system will have a positive and significant impact on its use.
H2b. The quality of the system will have a significant positive effect on user satisfaction.
H4. The intensity of use has a significant positive effect on the net results obtained.

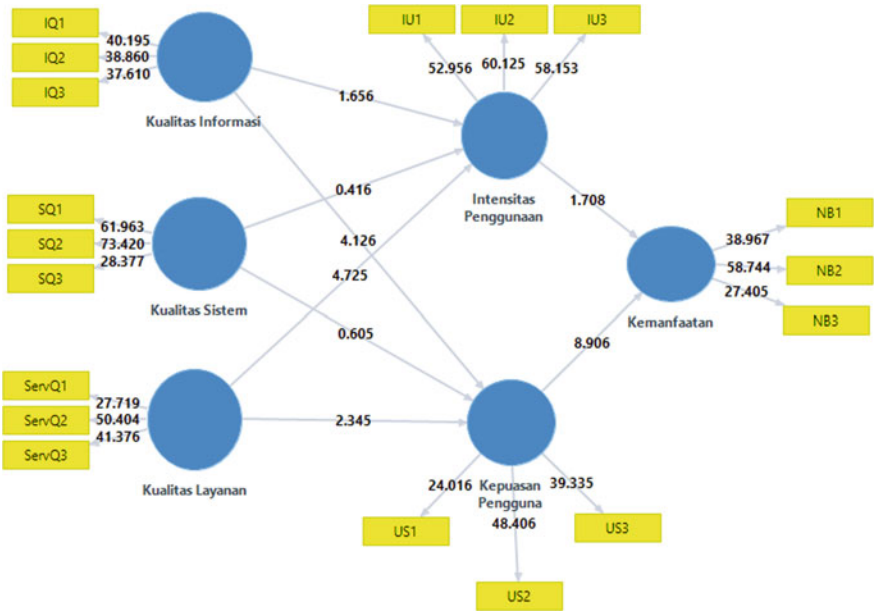


Fig. 2 Output bootstrapping path coefficient

Hypothesis 1b is supported by evidence indicating the quality of material on a Web site has a positive effect on user or community satisfaction. The higher the quality of the information on a Web site makes the happiest consumers. These results are similar with prior study [9], which examined the performance of the Wonosobo district government Web site and found that the independent variable, information quality had an effect on user happiness. This demonstrates that consumers are becoming happiest with the Web site's functionality as a result of the Web site's information quality. The similar conclusion was reached by [7], who observed the quality of information supporting user satisfaction in Kota Pariaman's village banking system. Thus, the findings of this study corroborate Delone and Mclean's theory that the quality of information on a Web site can affect community users' satisfaction with the Ngawi district government Web site. While hypothesis H1b is acceptable, hypothesis H1a demonstrates that information quality has no positive or substantial influence on how local government Web sites are used. As a result, this hypothesis is rejected due to its low t-statistic value. This notion is also refuted by the results of who evaluated the national CPNS selection system's Web site. Not only that but, this idea was also contradicted by the research, which observed the Darmo Surabaya hospital's Web site's success. Thus, the quality of the material shown on the Web site has no effect on how the public uses it, contradicting Delone and Mclean's argument.

The rejection of hypotheses H2a and H2b demonstrates that the system's quality has no positive and statistically significant influence on Web site users' usage and satisfaction. This outcome is consistent with the study undertaken by, which resulted

in the system quality hypothesis being rejected on the national CPNS selection system's measuring Web page. Not just on government Web sites but also in the assessment of e-commerce Web sites, this notion is rejected. According to research, the quality of the system has no effect on the happiness of customers of the e-commerce Web site JualdanBeli.com. As a result, the quality of the system in this research does not support Delone and Mclean's claim that it has an effect on the usage and happiness of information system users.

Meanwhile, the accepted hypotheses H3a and H3b indicate that the quality of service on the Web site has a positive and significant effect on the use and satisfaction of users in accessing the Web site. The better the quality of Web site services will increase the use and satisfaction of users. This finding is the same as research from [7] which shows the quality of service influences user satisfaction. Although the two hypotheses in this study were accepted, this research shows the quality of service does not affect the use of the Web site. Thus, the two hypotheses accepted in this study support the Delone and Mclean theory that service quality influences user use and satisfaction in accessing the Web site.

Hypothesis H4 demonstrates that the hypothesis is rejected, indicating that the user has no positive or substantial influence on the Web site's results/benefits. This notion is also refuted by the results of e-commerce Web sites. This conclusion is corroborated by research from, which assesses the Web site's success in Balikpapan. In contrast to study, the hypothesis is acknowledged in Brawijaya University's measurement of information systems. The accepted notion is further supported by the results of, who analyzed the Malang city's health information system. As a consequence, although the findings of this research do not support the Delone and Mclean theory, support for the idea may exist in other investigations.

Furthermore, hypothesis H5 shows that user satisfaction positively affects the results/benefits obtained from accessing the Web site. This finding is the same as research from which examines the success of e-commerce Web sites with the Delone and Mclean model. The same result is also obtained from, who saw user satisfaction on the net results obtained on the research and community service information system at Brawijaya University. The acceptance of this hypothesis is also strengthened by research on the health information system in Malang City. Thus, the satisfaction of the Ngawi Regency Government Web site users influences the net results/benefits obtained from accessing the Web site, thus supporting the theory of Delone and Mclean.

5 Conclusion

The Ngawi Regency Local Government Web site in providing public information is measured for its success using the Delone and Mclean model, which produces the following conclusions all indicators are declared valid, and the proposed variables have high reliability. There are eight hypotheses proposed and resulted in four accepted and four rejected hypotheses. The quality of information on the Web

site influences people's satisfaction in accessing local government Web sites. The quality of the system on the Web site does not influence the use and satisfaction of the community in accessing local government Web sites. The quality of service on the Web site influences the use and satisfaction of the community in accessing local government Web sites.

References

1. Cahyono TA, Susanto TD (2019) Acceptance factors and user design of mobile e-government website (Study case e-government website in Indonesia). Procedia Comput Sci 161:90–98
2. Cheisviyanny C, Helmy H, Dwita S (2018) Analisis Kualitas Website Pemerintah Daerah Kabupaten/Kota di Provinsi Sumatera Barat. Simp. Nas. Keuang. Negara 6(3):1087–1104
3. Yunita NP, Aprianto RD (2018) Kondisi Terkini Perkembangan Pelaksanaan E-Government di Indonesia. In: *Seminar Nasional Teknologi Informasi dan Komunikasi*, pp 329–336
4. Hernikawati D (2016) Analisis Popularitas Website Tingkat Kementerian. J. Masy. Telemat. dan Inf. 7(2):79–88
5. Masyhur F (2014) Kinerja Website Resmi Pemerintah Provinsi di Indonesia Official Website Performance Local Government in Indonesia. J. Pekommas 17(1):9–14
6. Muttaqin MH, Susanto TD (2019) The effect of website components on user trust in increasing the interest to use public administration service on e-government website. In: Proceedings of 2019 International conference on computer science, information technology, and electrical engineering ICOMITEE 2019, vol 1, pp 30–36
7. Dutta A, Devi MS, Arora M (2017) Census web service architecture for e-governance applications. In: ACM international conference proceeding series, vol Part F1280, pp 1–4
8. Cumbie BA, Kar B (2016) A study of local government website inclusiveness: the gap between e-government concept and practice. Inf Technol Dev 22(1):15–35
9. Jiang, Ji (2014) E-government web portal adoption: the effects of service quality. e-Service J 9(3):43
10. Henman P, Graham T (2018) Webportal vs google for finding government information on the web: from a website-centric approach to a web ecology perspective. Inf Polity 23(4):361–378
11. Delone WH, Mclean ER (2003) The DeLone and McLean model of information systems success: a ten-year update. J Manag Inf Syst 19(4):9–30
12. McGill T, Hobbs V, Klobas J (2003) User-developed applications and information systems success: a test of DeLone and McLean's model. Inf Resour Manag J 16(1):24–45
13. Delone WH, Mclean ER (1992) Information system success: the quest for the dependent variable. Inf Syst Res 3(1):60–95
14. Delone WH, McLean ER (2016) Information systems success measurement. Founda Trends Inform Syst 2(1)
15. Jung Y, Perez-Mira B, Wiley-Patton S (2009) Consumer adoption of mobile TV: examining psychological flow and media content. Comput Human Behav 25(1):123–129
16. Afnan A (2018) Pengukuran Keberhasilan Penggunaan E-Ticketing PT Kereta Api Indonesia (KAI) dengan Modified Delone Dclean. J. Bingkai Ekon. 3(2)

Factor Influencing Trust in Government: A Survey in the Bantul Regency

Denny Ardiansyah Pribadi, Ulung Pribadi, Dyah Mutiarin, and Vindhi Putri Pratiwi

Abstract This study aims to determine the factors that influence public trust in the government. Indonesian people's belief in the government has begun to decline, especially during the COVID-19 pandemic. So that it is crucial in this study to test the theory of the independent hypothesis built to see the relationship between people's intentions to trust the government. This study uses data from questionnaires that the people of Bantul Regency have given responses to. The number of research respondents was 100 respondents and processed using SmartPLS 3.0 Software. The results showed that of the three hypotheses tested; all showed a positive relationship concerning trust in the government. Variable perceived usefulness has become an accepted variable according to the hypothesis. This research is used as input for the Bantul Regency Government to increase public trust in the government, mainly focusing on variables that have positive implications.

Keywords Trust · Government · Society

1 Introduction

The increasing dynamics of the flow of information, communication, and technology take society in an important position. Social patterns in society have changed. Not only is community today subject to government power and policies but above all, companies are in a very strategic position to define a growing continuum of activity, government, and other social organizations. One of the problems between the government and the community is the trust of the government. As a result, the issue of crisis and loss of trust is increasing concern by sociologists and is the subject of current research [1]. Government trust is an essential thing in the implementation of government activities. This trust is related to the public's belief in the services provided by

D. A. Pribadi (✉) · U. Pribadi · D. Mutiarin · V. P. Pratiwi
Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: denny.a.psc20@mail.umy.ac.id

V. P. Pratiwi
e-mail: vindhi.putri.psc19@mail.umy.ac.id

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_4

the government. The more assistance offered by the government run well, the more people will trust the government. However, if the services offered by the government are not going well, the public will lose trust in the government.

Previous research has discussed a lot about trust in government. The difference between this study and prior studies lies in the direct discussion on the implementation and the research locus. As a part of government activities, using research variables, this study examines the elements that determine trust in the government in Bantul Regency. Perceived ease of use, perceived usefulness, and security risk as dependent variables, and trust in government as an independent variable. This study aims to determine and describe the level of public trust in the Bantul Regency Government. This research expects to be an evaluation material for the State Civil Apparatus in the Bantul Regency Government to improve their ability to serve the community to build public trust in the government.

2 Literature Review

2.1 Trust in Government

Trust is an essential point in the implementation of good governance. Trust also defines that attitudes influence a person's behavior in the form of belief in the circumstances around the individual [2]. Trust here means the relationship between the individual and the organization. The surrounding circumstances influence the relationship between the individual and the organization, so trust is firmly bound between the individual and the organization. Trust in the government is essential in the implementation of government activities. The level of public trust in the government has a relationship with community participation related to the government program. It means that the higher the public participation in government programs, the higher the level of trust in the government. Conversely, if public involvement related to government programs is low, then the level of confidence in the government is lacking.

2.2 Perceived Ease of Use

Perceived ease of use means the extent to which a person believes that technology frees him from the effort. Perceived ease of use also refers to the amount to which a person's level of trust in a system can increase his work performance [3]. Perceived ease of use here has a meaning related to one's belief in using technology. Trust in technology is expected to increase one's ability to improve performance so that the ability to complete work is good. Technology in perceived ease of use is more

used in various things, such as e-mail, e-commerce, and applications needed by the government [4].

2.3 Perceived Usefulness

Perceived usefulness means the degree to which a person in increasing his work capacity has more confidence in using a particular system [5]. Perceived usefulness determines how a person's confident attitude toward the use of innovation [6]. Perceived usefulness here has a meaning related to the level of trust in technology. Because in this day and age, technology has become essential in all lines of work. So, every job combined with technology will provide convenience to its users.

2.4 Security Risk

Technical, organizational, scientific, legal, economic, and information challenges all pose security risks in e-government [7]. This area is addressed through security mechanisms, careful system design, and policy measures [8]. Security mechanisms are created with the aim of detecting, preventing, and recovering from security attacks. It is usually using several approaches, such as cryptographic firewalls and information system security. The installation and configuration of hardware, software, and data recovery (backup) are all part of system design to ensure that the e-government system runs properly. Administrators and authorities implement policies (including assessments) by establishing procedures for planning, testing, certifying, monitoring, evaluating, and certifying e-government systems. Another model, called the e-government maturity model (e-GMM), guides and compared the development of e-government from a regulatory perspective. For example, Karokola et al. presented the e-government maturity model (e-GMM) to guide the integration of security in e-government systems with: policymakers in the government [9].

3 Methods

This study uses a quantitative approach. The quantitative method aims to use data and numbers as a tool to gain knowledge, analyze what you want to know (), and describe the data collected. This survey data collection method uses a survey method using a survey tool—researchers surveyed by distributing questionnaires to residents of Bantul Regency. Using a sampling error of 10%, where the population has the same population characteristics, this number of samples is 99.98 or as many as 100 respondents. Data processing uses the PLS-SEM method, which is to perform multiple regression. Data analysis uses descriptive quantitative techniques in the

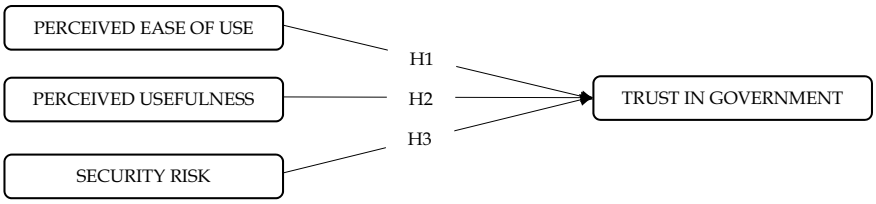


Fig. 1 Research hypothesis (processed by researchers)

form of factor analysis that affects trust in government. Another hypothesis validity test in this study combines them using the PLS structural model, which evaluates using the R^2 of the dependent structure, the path coefficient value, or each path's t value. The following describes the hypothesis used in this study.

Based on the explanation and theoretical framework in Fig. 1, the researcher makes the following hypothesis:

H1. Perceived ease of use has positive and significant implications for trust in government.
H2. Perceived usefulness has positive and significant implications for trust in government.
H3. Security risk has positive and significant implications for trust in government.

4 Result and Discussion

Based on a survey conducted by researchers, data were obtained from 100 respondents in Bantul Regency. The respondents were divided into 53 men and 47 women. Furthermore, the majority of respondents are between 36 and 45 years old. As for the composition of the last education level taken, most of the respondents obtained by the researcher came from respondents who are currently/have completed their studies at the lecture level (Table 1).

4.1 Validated Research Model, Data Validity, and Data Reliability

The data testing model used by the researcher perceived ease of use, perceived utility, security risk, and privacy risk, all of which influence faith in government or act as a mediating variable in affecting the intention to utilize e-government. The quality of the measurement model is seen from the validity and reliability by considering several values, namely: convergent and discriminant validity are two types of validity. These two numbers can be seen in the average variance extracted score results from the outer model (AVE). The level of data validity can be said to be valid if the AVE

Table 1 Profile of respondents (processed by researchers)

Legends	Frequency	%
Gender		
Men	53	53
Woman	47	47
Age		
17–25	28	28
26–35	20	20
36–45	36	36
46–55	16	16
56–65	0	0
Level of education		
Primary school	0	0
Junior high school	3	3
Senior high school	15	15
Associate's degree	18	18
Bachelor	53	53
Master	11	11

score is more than 0.5. However, if it is below 0.5, then the AVE score cannot be said to be valid [10, 11] (Fig. 2).

From the results of the validity test in Table 2, we can see that all statements on the variables are as follows: perceived ease of use, perceived usefulness, security risk, and trust in government. These variables have outer loading values >0.5, and

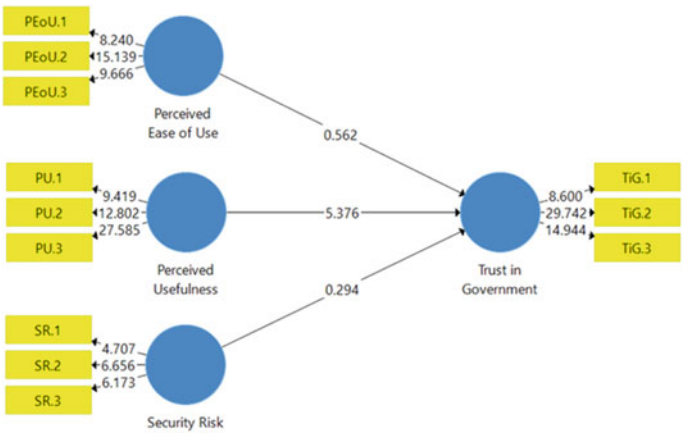


Fig. 2 Validated research model (processed by researchers using SmartPLS 3.0)

Table 2 Data validity (processed by researchers using SmartPLS 3.0)

Variable	Indicator	Outer loadings	Average variance extracted	Description
Perceived ease of use	PEoU.1	0.814	0.677	Valid
	PEoU.2	0.855		
	PEoU.3	0.798		
Perceived usefulness	PU.1	0.722	0.640	Valid
	PU.2	0.797		
	PU.3	0.873		
Security risk	SR.1	0.765	0.662	Valid
	SR.2	0.863		
	SR.3	0.810		
Trust in government	TiG.1	0.742	0.658	Valid
	TiG.2	0.878		
	TiG.3	0.807		

all of them have AVE values >0.5. Thus, the authors conclude that the questions in all variables are valid or meet convergent validity and discriminant validity.

The mean value in Table 3 shows the degree of correspondence between the three independent variables and one dependent variable. The independent variable with the highest value is "Perceived ease of use," with a value of 4.307, followed by "perceived usefulness" of 4.190. This variable is a common problem in Indonesia. The perception of use and ease of application are often common problems in Indonesia because there are so many applications that the government presents but the lack of socialization and lack of readiness from the community in welcoming the technology [12].

Table 3 also shows Cronbach's alpha and composite reliability values >0.6 and 0.7, respectively. So, we can conclude that all research variables, both independent variables (perceived ease of use, perceived usefulness, security risk) and dependent variables (trust in government), meet the reliability test. After the reliability test, we will measure the inner model.

Table 3 Data reliability (processed by researchers using SmartPLS 3.0 software)

Variable	Mean	Standard deviation	Cronbach's alpha	Composite reliability	Description
Perceived ease of use	4.307	0.589	**0.762**	**0.862**	Reliable
Perceived usefulness	4.190	0.685	**0.717**	**0.841**	Reliable
Security risk	3.950	0.699	**0.751**	**0.855**	Reliable
Trust in government	4.107	0.686	**0.738**	**0.850**	Reliable

Table 4 Number of R-squares (processed by researchers using SmartPLS 3.0)

Variable	R-square
Trust in government	0.301

4.2 Inner Model Test

Researches will test the inner model test to see the relationship between the constructs, the significance value, and *R*-square based on the research model. The significance of the coefficients of the structural route parameters and the R-square for the dependent construct of the t-test will be used to evaluate the structural model. In assessing the model with PLS, it starts by looking at the R-square for each latent-dependent variable [11]. Table 4 is the result of calculating the R-square with SmartPLS 3.0.

According to Hamid and Anwar [13], when connected with the rule of thumb, the R-square test is 0.75 in the strong category, 0.50 in the medium category, and 0.25 in the weak category [13]; the authors conclude that the variables that affect trust in government have an influence value in the moderate category. Its shows that 30% of the trust in government variables is influenced by perceived ease of use (PEoU), perceived usefulness (PU), and security risk (SR).

4.3 Hypothesis Testing

After the data gathered were valid and reliable, the bootstrap resampling approach was used to test hypotheses between variables, specifically exogenous variables to endogenous variables and vice versa. The T statistic is the test statistic employed (t-test). The *t* value for comparison is taken from the t table. The test can be significant if the T statistic value >1.96 and *P*-value <0.05. The hypothesis testing may be seen in Table 5 by looking at the output path coefficient from the bootstrap resampling findings.

Table 5 shows that of the three variables, only one hypothesis is accepted. The accepted hypothesis is marked with a T statistic value >1.96 and a *P*-value <0.05 (green). At the same time, the rejected hypothesis means that it does not meet the T statistic value >1.96 and *P*-value <0.05 (red color).

Table 5 Hypothesis testing (processed by researchers with SmartPLS 3.0).

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	P Values	Result
PEoU->TiG	-0,052	-0,039	0,093	0,562	**0,574**	Rejected
PU->TiG	0,562	0,552	0,104	5,376	**0,000**	Accepted
SR ->TiG	0,035	0,061	0,118	0,294	**0,769**	Rejected

H1. Perceived ease of use has no significant implications for rust in government (the hypothesis is rejected), with a T statistic of 0.562 and a *P*-value of 0.574. This study also confirms the findings of previous studies that perceived ease of use has a significant effect on trust in government [14–16]. This result looks surprising because a problem that often occurs in Indonesia is the ease of use.

H2. Perceived usefulness has significant implications for trust in government (the hypothesis is accepted), with a T statistic of 5.376 and a *P*-value of 0.000. This study confirms the results of previous studies that perceived usefulness has a significant effect on trust in government [15, 16].

H3. Security risk has no significant implications for trust in government (the hypothesis is rejected), with a T statistic of 0.294 and a *P*-value of 0.769. This study also confirms the findings of previous studies that security risk has a significant effect on trust in government [15, 17]. It seems surprising because security issues are the subject of discussion in Indonesia.

5 Conclusion

This study contributes to empirically testing the perceived ease of use, perceived usefulness, and security risk on trust in government conducted by the Bantul Regency Government. Because the relationship between variables has not been widely studied, this study can be used as one way to increase the level of public trust in the government in Indonesia.

There is only one variable with positive implications and significant value. The variable is in the form of perceived usefulness of trust in government. Meanwhile, two variables have negative implications and have no significant value. The variables are perceived ease of use against trust in government and security risk against trust in government.

The practical implications of this research can be used as input for the Bantul Regency Government to increase public trust in the government, mainly focusing on variables that have positive implications. This research also still has obstacles in the form of a small research locus, a small number of respondents, a short research time, and research variables. Thus, it is hoped that this paper will become a recommendation for further research.

References

1. Silalahi U (2011) Kepercayaan Publik kepada Pemerintah Daerah Pasca Orde Baru. Ilmu Akunt. Negara 11(2):167–179
2. Yulianti A, Kurniawan A (2019) Pengaruh Pengetahuan Perpajakan Dan Persepsi Keadilan Terhadap Tingkat Kepatuhan Wajib Pajak Melalui Kepercayaan. J. Ris. Keuang. Dan Akunt. 5(1). https://doi.org/10.25134/jrka.v5i1.1929

3. Hamid AA, Razak FZA, Bakar AA, Abdullah WSW (2015) The effects of perceived usefulness and perceived ease of use on continuance intention to use E-Government. Procedia Econ Financ 35:644–649. https://doi.org/10.1016/s2212-5671(16)00079-4
4. Moslehpour M, Pham VK, Wong, WK, Bilgiçli I (2018) e-purchase intention of Taiwanese consumers: sustainable mediation of perceived usefulness and perceived ease of use. Sustainability 10(1). https://doi.org/10.3390/su10010234
5. Alsabawy AY, Cater-Steel A, Soar J (2016) Determinants of perceived usefulness of e-learning systems. Comput Human Behav 64:843–858. https://doi.org/10.1016/j.chb.2016.07.065
6. Sugandini D et al (2018) The role of uncertainty, perceived ease of use, and perceived usefulness towards the technology adoption. Int J Civ Eng Technol 9(4):660–669
7. You C (2020) Law and policy of platform economy in China. Comput Law Secur Rev 39:105493. https://doi.org/10.1016/j.clsr.2020.105493
8. Al-Rodhan MA, Janabi, Sekhan (2011) Isolation and identification of *Staphylococcus* spp. from Bovine Mastitic milk and their sensitivity to some antibiotics at Al-Qadissiyah Province. J Vet Med Sci 10(2)
9. Mukred M, Yusof ZM, Alotaibi FM, Mokhtar UA, Fauzi F (2019) The key factors in adopting an electronic records management system (ERMS) in the educational sector: a UTAUT-based framework. IEEE Access 7:35963–35980. https://doi.org/10.1109/ACCESS.2019.2904617
10. Kelibay I, Nurmandi A, Malawani AD (2020) e-government adoption of human resource management in Sorong City, Indonesia. *J*. Asian Rev. 5(1):1–20. [Online]. Available https://arpap.kku.ac.th/index.php/arpap/article/view/180.
11. Kusuma MDP, Pribadi U (2020) Faktor-Faktor Yang Mempengaruhi Niat Perilaku Pengguna Aplikasi 'Jogja Istimewa' Melalui Model Umega. Moderat J. Ilm. Ilmu Pemerintah. 6:771–791. [Online]. Available https://jurnal.unigal.ac.id/index.php/moderat/article/view/4031.
12. Margono M, Masithoh RF, Priyo P (2020) Preparedness prevention of Covid-19 Muhammadiyah residents in Magelang, Indonesia. J. Ilm. Kesehat. Keperawatan 16(2):01. https://doi.org/10.26753/jikk.v16i2.492
13. Hamid RS, Anwar SM (2019) Structural equation modeling (SEM) Berbasis Varian. PT. Inkubator Penulis Indonesia, Jakarta
14. Mujiasih R, Wiwoho G (2020) Pengaruh Perceived Usefulness, Perceived Ease of Use Terhadap Trust dan Intention to buy KAI Access Tickets Online pada Generasi Millennial di Kabupaten Kebumen. J. Ilm. Mhs. Manajemen, Bisnis dan Akunt. 2(3):476–490. https://doi.org/10.32639/jimmba.v2i3.492
15. Nangin MA, Barus IRG, Wahyoedi S (2020) The effects of perceived ease of use, security, and promotion on trust and its implications on fintech adoption. J Consum Sci 5(2):124–138. https://doi.org/10.29244/jcs.5.2.124-138
16. Sito Putri RR, Iriani SS (2021) Pengaruh Perceived Ease of Use dan Perceived Usefulness terhadap Keputusan Penggunaan Aplikasi Tokopedia melalui Trust sebagai Variabel Intervening. J. Ilmu Manaj. 9(2):708. https://doi.org/10.26740/jim.v9n2.p708-722
17. Afiah N (2018) Pengaruh Keamanan, Reputasi Dan Pengalaman Terhadap Trust Pengguna Internet Untuk Bertransaksi Secara Online. JEKPEND J. Ekon. dan Pendidik. 1(2):58. https://doi.org/10.26858/jekpend.v1i2.7256
18. Jeewa A, Kasiram M (2008) Treatment for substance abuse in the 21st century: a South African perspective. South Afr Fam Pract 50(6):44. https://doi.org/10.1080/20786204.2008.10873782

Different Applications and Technologies of Internet of Things (IoT)

Feisal Hadi Masmali, Shah J. Miah, and Nasimul Noman

Abstract Internet of things (IoT) has significantly altered the traditional lifestyle to a highly technologically advanced society. Some of the significant transformations that have been achieved through IoT are smart homes, smart transportation, smart city, and control of pollution. A considerable number of studies have been conducted and continues to be done to increase the use of technology through IoT. Furthermore, the research about IoT has not been done fully in improving the application of technology through IoT. Besides, IoT experiences several problems that need to be considered in order to get the full capability of IoT in changing society. This research paper addresses the key applications of IoT, the architecture of IoT, and the key issues affecting IoT. In addition, the paper highlights how big data analytics is essential in improving the effectiveness of IoT in various applications within society.

Keywords Internet of things (IoT) · Healthcare service · Information technologies · IT adoption · Big data analytics

F. Hadi Masmali (✉)
School of Electrical Engineering and Computing, The University of Newcastle, Newcastle, New South Wales, Australia
e-mail: feisal.masmali@uon.edu.au

College of Business Administration, Jazan University, Jizan, Saudi Arabia

S. J. Miah
Newcastle Business School, The University of Newcastle, Newcastle City Campus, Newcastle, New South Wales, Australia
e-mail: shah.miah@newcastle.edu.au

N. Noman
School of Information and Physical Sciences, University of Newcastle, Newcastle, NSW, Australia
e-mail: nasimul.noman@newcastle.edu.au

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_5

1 Introduction

The term "Internet of Things" (IoT) is used to illustrate a network of physical objects that is, "things" that are implanted with software, sensors as well as other technologies to facilitate connectivity as well as the exchange of information with other systems as well as devices over the Internet [1]. Such devices range from objects in households to complex industrial tools. According to Rangers [1], there are above 7 billion connected IoT gadgets in the world, and the number is approximated to rise to 22 billion by 2025. As an upcoming paradigm, IoT allows communication between electronic devices and sensors via the Internet, which significantly facilitates people's lives. In addition, IoT applies smart devices as well as the Internet to develop solutions to challenges as well as issues that various businesses and private as well as public industries face across the world [2]. Furthermore, IoT has continuously turned out to be a vital aspect of our lives, which is evident everywhere in the world. In general, IoT integrates different varieties of smart systems frameworks as well as intelligent devices as well as sensors [2]. Besides, IoT exploits quantum as well as nanotechnology in the form of sensing, storage as well as processing speed, which was not attainable earlier [3]. There have been extensive studies that have been conducted in order to illustrate the effectiveness as well as the applicability of IoT transformations. This paper provides an in-depth illustration of IoT in terms of its architecture, its major applications, significant problems and issues of IoT, and the significance of big data analytics within IoT.

There is a significant transformation that is evident from our daily lives whereby the application of IoT devices as well as technology has significantly increased. The establishment of Smart Home Systems (SHS), home automation systems, and appliances that use devices that rely on the Internet are some of IoT developments [4]. Another significant gain of IoT is the Smart Health Sensing System (SHSS). This system integrates several small intelligent pieces of equipment as well as devices that assist in supporting health. The system is also applied to inspect critical health conditions within the hospitals. This signifies that IoT has altered the medical domain scenario by introducing convenient technology as well as smart devices [5]. Furthermore, the developers, as well as researchers of IoT, have actively been participating in improving the lifestyle of both disabled and aged individuals, whereby IoT has an extreme performance within this sector. IoT has indicated a new direction for the ordinary life of the aged and disabled individuals through the development of devices that are cost-effective and easy to use for these individuals [2].

The other important aspect of IoT in daily life is transportation. It has contributed to new advancements making transportation more efficient, comfortable as well as reliable. For instance, intelligent sensors and drone devices have been widely applied in controlling traffic across major cities. Additionally, there are new vehicles that consist of sensing devices that allow sensing heavy traffic congestions and to find alternative routes on the map with low traffic congestion [6]. This clearly indicates that IoT has played a significant role in different aspects of life as well as technology.

Besides, IoT has displayed its significance as well as potential within the economic and industrial fields. It has been regarded as a revolutionary tread within the trade and stock exchange market. Despite the potential associated with IoT in life improvement, the issue of security of information is a crucial concern, which is the main challenge faced in IoT [5]. Hackers have actively been using the Internet in conducting cyber-attacks, which has made data insecure. Despite this, IoT developers and researchers have been involved in giving the best possible solutions to cope with data security issues.

2 The IoT System Architecture

The architecture of an IoT system is typically illustrated as a process that involves steps whereby information flows from sensors incorporated in "things" via a network, and finally onto a corporate data center or the cloud whereby it undergoes processing, analysis as well as storage [7]. In this case, a "thing" may be a machine, a person, or a building. In the IoT system process, also send information to the other direction instructing the actuator or the connected devices to take various actions in order to control a physical process.

There are four essential layers incorporated in an IoT architecture that describe the workability of an IoT system. They are the perception layer, network layer, application layer, and business layer [2]. The perception is the lowest in the IoT structure. It constitutes physical devices such as sensors, Radio-frequency identification chips and barcodes, which are linked to the IoT system. They assist in gathering data that is then delivered to the network layer.

Moreover, the network layer functions as a transmission medium of data from the perception layer to the data processing system [2]. This channeling of data may apply either be through a wired or a wires medium. The application layer utilizes the processed data for global device management [2]. The business layer is found on the top of the IoT architecture. It usually controls the entire IoT system that is, its applications as well as services [22]. It visualizes data from the application layer and uses the information to generate plans as well as strategies. According to Kumar et al. [], the IoT architecture can be altered according to the need as well as the application domain (Fig. 1).

Apart from layered frameworks, IoT systems contain other functional blocks that help in various IoT services like sensing methods, authentication as well as identification, control, and management [8]. Besides, these functional blocks are involved in connectivity issues, audio or video monitoring, input or output operation, and management of storage. These blocks all form an effective IoT system essential for maximum performance. According to Nicolescu et al. [9], several architecture references with technical specifications that have been proposed still deviate from the typical IoT architecture accepted for global IoT (Fig. 2).

According to Kumar et al. [2], the critical design factors for a productive IoT architecture within a heterogeneous condition are scalability, openness, modularity,

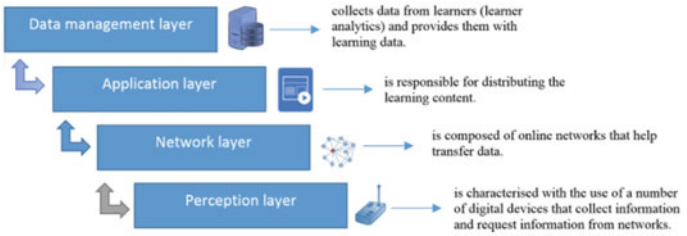


Fig. 1 The architecture of IoT system. *Source* Jahnke [7]

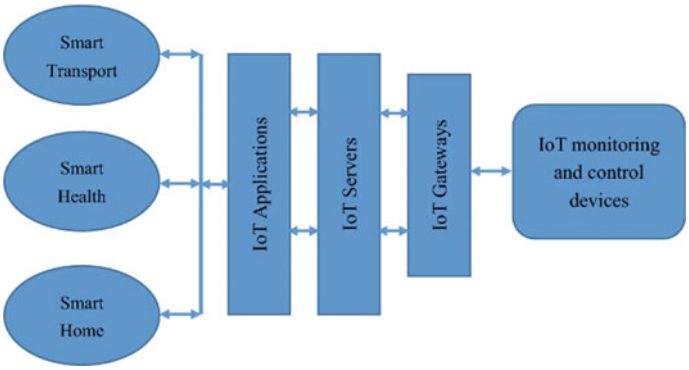


Fig. 2 The generic structure of IoT. *Source* Kumar et al. [2]

and inter-operability. The design must include an objective of meeting the needs of cross-domain interactions, big data analytics, storage and applications that are user-friendly. The IoT structure must also have the ability to scale up the workability and include intelligence as well as automation within the IoT devices [2].

Furthermore, the increased amount of data generated from IoT sensors and IoT gadgets has contributed to a big issue. This implies that an effective IoT structure should be able to cope with the increased amount of data within the IoT system. The leading IoT system architectures are fog and cloud computing. They are responsible for handling, monitoring as well as analysis of a large amount of information within IoT systems [2].

Sensors, as well as actuators, play a vital role in stage one of the IoT architecture. Sensors have the capacity to detect data flow from the real world, which constitutes things like humans, vehicles, animals, electronic devices, and buildings. The sensors transmit this data, which may be used further for analysis. On the other hand, actuators are capable of intervening the reality, for instance, turning off the music, slowing down the speed of a vehicle, and regulating temperature in a room. In general, stage one of the IoT architecture enables collecting information from the real world. This information may be applied for more analysis.

Moreover, stage two enables cooperation between sensors and actuators together with gateways as well as information acquisition systems. In addition, this stage allows aggregation and optimization of the enormous data gathered from stage one to appropriate way for the purpose of processing [2]. The data is then passed to stage three of the IoT architecture, which is edge computing. This sage can be described as a free architecture that enables the utilization of technologies as well as huge computing power from different areas in the world [2]. In addition, edge computing provides a robust way of streaming data processing, and therefore it is acceptable for IoT systems.

During stage three of IoT architecture, technologies of edge computing manage a huge amount of information and give several functionalities like imaging, data integration from different sources, and analysis of data through machine learning techniques. The final stage involves various crucial activities like deep analysis and processing of information, giving feedback to increase the whole system's precision. All processes in this stage are conducted in the cloud server [2]. Moreover, this stage involves big data frameworks like Spark in order to handle substantial streaming data. Furthermore, machine learning methods may be applied to establish suitable prediction models that may help develop a reliable and accurate IoT system that meets the requirements of people [2].

3 Major Applications of IoT

IoT has a wide range of applications since it can be adjusted to almost every technology that can provide suitable information concerning its operation, the fulfillment of an activity, and concerning the environmental situations that require to be monitored and regulated distantly. A significant number of industries are currently adopting IoT technologies to simplify, enhance, and automate as well as to control distinct processes.

4 Smart City, Transport as Well as Vehicles

IoT has significantly altered the conventional civil structure of society to form a high technological structure that involves concepts of smart city, smart home as well as smart vehicles. There are vast advancements that are aided by supporting technologies like machine learning [10]. Furthermore, several technologies like cloud server technology as well as wireless sensor networks and IoT servers must be applied jointly in order to generate an efficient smart city. However, it is vital to consider the environmental factor of the smart city. In connection to this, there should be the establishment of energy-efficient technologies as well as green technologies when it comes to the establishment of smart city infrastructure.

Furthermore, new vehicles are currently incorporated with smart devices, which can easily detect traffic congestion on roads and helps in providing alternative routes. This significantly helps in minimizing congestions n within the city. Besides, there should be a design of smart devices with optimum cost that can be compatible with all vehicles in monitoring different parts of the vehicle. In addition, IoT significantly helps in maintaining the condition of a vehicle. Also, self-driving vehicles have the capacity to pass information to other vehicles via intelligent sensors. This allows smooth flow of the traffic compared to human-driven vehicles. Although this technology will require significant time to spread worldwide, IoT gadgets will continue helping in sensing traffic congestion.

Sensors in fleet vehicles facilitate effective communication between the managers of the vehicle and the vehicle. The manager or the driver of the vehicle can easily get information concerning the status, operation as well as requirement of the vehicle by evaluating the software responsible for gathering, processing, and organizing information. In general, the integration of IoT to fleet management helps in geolocation and performance analysis and may allow the generation of relevant information that can help improve driving vehicles.

5 Maintenance Management

This forms one of the extensive deployment areas of IoT technology. The amalgamation of sensors and software specialized in Enterprise asset management (EAM) as well as computerized maintenance management systems (CMMS) allowed the development of a multifunctional tool that can be applied in various disciplines as well as practices. The main objective of these tools is to increase the productive life of assets. The application of these tools is almost unlimited since the software attributes that process and organize data gathered by sensors are specifically designed to address the management requirements of physical assets [11].

6 Healthcare

According to Kumar et al. [2], IoT is entirely committed to providing financial benefits as well as development to society. These developments include but are not limited to public facilities such as water and quality maintenance, economic development, and industrialization. When it comes to the health sector, the application of wearable sensors that are linked to patients enables healthcare providers to observe a patient's condition while outside the healthcare facility or in real-time. Via progressive inspection of various metrics as well as automatic alerts about their crucial signs, the IoT assists in enhancing patients care and prevention of lethal situations among high-risk patients [11]. Moreover, the integration of IoT in hospital beds is another essential application of IoT within the healthcare sector. This has given way to smart beds that

are incorporated with sensors to monitor patients' vital signs such as blood pressure and body temperature.

7 Wearables Devices

These are small devices that do not consume a lot of energy and are equipped with sensor and are equipped with the relevant hardware to take measurements as well as readings and their software gather and organize data concerning the users. Some of the examples of wearable devices that have been widely applied are virtual glasses, GPS tracking belts, and fitness bands to observe, for instance, calories use. Corporations like Google and Samsung have established and are using IoT in their devices [11].

8 Hospitality

IoT integration within the hotel industry enables improvement in the provision of services. For instance, the implementation of electronic keys that are directly sent into guests' mobile devices allows the automation of several interactions. Some of the crucial activities that can be easily managed through the application of IoT in the hotel industry are the realization of orders or services in a room and the automated charging of room [11]. Moreover, the process of checking out is automated with the help of electronic keys.

9 Smart Grid as Well as Energy Saving

IoT has enabled better monitoring as well as control of the electrical network. This has been effectively achieved by the continuous application of smart meters with sensors, and installing sensors at various strategic locations from the point of power production to distribution areas. These sensors significantly allow bidirectional communication between the power companies and users, whereby essential information pertaining to faults, repairs, or decision-making can be obtained. Moreover, IoT allows sharing critical information to the end-users concerning their power consumption patterns and concerning the best way of reducing their expenditure on power [11].

10 Water Supply

Moreover, IoT has progressively been applied in the supply of water. A sensor that is equipped internally or externally in water meters, connected to the Internet, and

integrated with the relevant software assist in collecting, processing, and analyzing data. This information is essential as it allows detecting faults, understanding water consumers' behaviors, reporting the outcomes, and providing a course of action to the company. Similarly, the application of IoT in water supply allows the consumers to track information about their water consumption via a webpage and even in real-time, whereby they can get alerts when consumption is exceeded the average consumption.

11 Agriculture

Agriculture is a vital sector to the rapidly growing world's population. Therefore, there is a need for the advancement of present agricultural activities in order to satisfy the huge world's population. This calls for the amalgamation of agriculture with advanced technology in order to increase production levels. One of the probable methods is through greenhouse technology toward this direction. Greenhouses provide a way of controlling environmental parameters to increase production. However, this has been improved by automating processes through the application of IoT. Through smart devices as well as sensors, environmental parameters can be easily monitored. This contributes to improved production and energy saving.

12 Key Issues Associated with IoT

There are several issues associated with the inclusion of IoT in almost all aspects of daily life and technologies associated with the transmission of data between the IoT integrated devices. Besides, these issues have raised challenges to IoT developers and researchers with the advanced smart society. Moreover, as technology continues to progress, there is also a need to advance IoT systems.

13 Privacy and Security

One challenge facing IoT systems is security as well as privacy. This is due to the available threats, cyber-attacks as well as vulnerabilities [12]. Some of the factors that contribute to privacy issues on IoT embedded devices include inadequate authorization as well as authentication, web interface, firmware, insecure software, and faulty transport layer encryption [13]. According to Sisinni et al. [14], security, as well as privacy, are essential in confidence development in IoT systems. In order to prevent attacks, security techniques should be incorporated at each layer of IoT [15]. Furthermore, various protocols have been established at each layer of communication in order to ensure that IoT-based systems are secure from threats [16]. An example of a security protocol that has been established between the transport and application

layer in order to secure IoT systems is the Secure Socket Layer (SSL). Different IoT systems require varied methods in the provision of communication security solutions between IoT devices. Furthermore, wireless communication between IoT devices increases the risk of security issues.

On the other hand, privacy is a vital aspect of IoT as ill allows IoT devices users to feel safe while applying IoT systems. It is critical to maintaining authorization as well as authentication in a safe network to develop trust between parties communicating with each other [17]. Moreover, there are various privacy policies for distinct devices within the IoT systems. To counter this issue, each object should have the capacity to verify other objects' privacy policies within the IoT system before transferring information.

14 Inter-Operation and Standard Issues

Inter-operability or integration refers to the practicability in the switching of data among several IoT devices as well as systems. Transmission of information does not necessarily depend on the devices and systems applied. Inter-operability issues emerge because of the heterogeneous nature of various technologies as well as solutions involved in the development of IoT. According to Inter-operability in Digital Health [18], the levels of inter-operability are technical, semantic-syntactic as well as organizational. Moreover, IoT systems incorporate several functionalities in order to improve inter-operability, which improves communication among various IoT objects within a heterogeneous environment. Some of the suggested ways of handling inter-operability are adapters/gateways-based and virtual networks/overlay-based [19].

15 Ethics, Law, and Rights

Ethics, legality, as well as regulatory rights are significant considerations for IoT developers. To safeguard morals and prevent people from breaching them, some specific policies and procedures should be considered. The sole difference that lies between ethics and legislation is that ethics are the guidelines that individuals are required to stick to, whereas legislations are the limitations pressed by the government. Moreover, ethics, as well as legislation, are intended to protect the quality and protect individuals from taking part in criminal behavior. Tzafestas [20] indicates that there are various real world problems, which have been resolved due to the development of IoT, but at the same time, resulting in various ethical as well as legal problems. These problems include privacy, security of information, trust, and safety of data usability. Due insufficient trust with IoT, many users align with government law regarding protection, privacy, and information security. Therefore, it is vital to

consider such issues in order to maintain and build up public trust in the deployment of IoT gadgets.

16 Dependability, Scalability, and Accessibility

When other devices are included in a system without decreasing its effectiveness, the system is regarded as scalable. The fundamental drawback with IoT is incorporating many devices with varying memory, processor, storage capability, bandwidth accessibility, and reliability [21]. Access to resources forms another significant factor that should be considered. Within the IoT architecture layers, scalability, as well as availability, should be applied jointly. A good example of a system that portrays scalability is the cloud-based IoT system. This is because they allow users to develop their IoT network by introducing other gadgets, storage, as well as processing capabilities as required.

The other important issue is making sure that resources to the suitable objects are accessible without necessarily bothering their location or time they are required. Several IoT networks are linked to global IoT systems in a scattered approach in order to maximize usage of their resources as well as services [2]. Various services, as well as resource availability, may be obstructed because of the availability of other different information transmission channels, like satellite communication. Therefore, an independent, as well as a dependable data transmission channel is required in order to ensure that resources and services are always available.

17 Quality of Service (QoS)

The other key issue in IoT systems is the quality of service (QoS) provided. QoS evaluates the quality, effectiveness, and productivity of IoT architecture, devices, and systems [22]. According to Huo and Wang [23], reliability, affordability, energy utilization, security, availability, and service time are significant in IoT systems. A good IoT platform should meet the necessities of QoS standards. Additionally, any IoT service should initially be defined to verify its dependability. People may also be free to articulate their requirements and needs accordingly. According to White et al. [24], many techniques can be used to analyze QoS, but a trade-off exists between quality variables and approaches. In order to address this trade-off, high-quality models must be used.

A few high-quality models like OASIS-WSQM may be applied in assessing the methodologies used for QoS evaluation [25]. Such models include a large number of quality parameters that are enough for determining QoS for IoT services.

18 Big Data Analytics in IoT

In an IoT system, many devices and sensors share information. In addition, the increased development of IoT networks has contributed to the increase of IoT devices and sensors at a rapid rate. There is, therefore, the transmission of an enormous amount of data by these sensors and devices through the Internet. The data transmitted is massive and, therefore, can be regarded as big data. Moreover, the increase in IoT embedded networks has resulted in issues like management and gathering of information, storage as well as processing, and analytics [2].

In smart buildings, the IoT big data system assists in dealing with issues like oxygen level management, detection of hazardous gases, and luminosity [26]. These systems enable the collection of information from sensors within the building and conduct data analytics to come up with decisions. Furthermore, Lee et al. [27], indicates that production in an industry can be raised by deploying an IoT embedded Cyber-physical platform that incorporates data analysis and knowledge gathering techniques.

Besides, traffic congestion within smart cities is a vital issue when it comes to big data analytics. Traffic light signals are able to gather real-time traffic data with the help of sensors and IoT devices incorporated in them. The information gathered can be easily analyzed within an IoT-based traffic management system [2].

Moreover, in the healthcare sector, IoT sensors help in generating more information concerning patient health. The huge amount of data gathered requires to be combined into one database for quick processing in order to enable developing accurate decisions. Big data technology effectively provides a solution for this [28]. According to Mourtzis et al. [29], when IoT is combined with big data analytics, it can effectively assist in changing the traditional methods applied in manufacturing industries. Devices that use sensors produce data that can be analyzed via big data techniques, which aid in making various decisions.

Besides, the application of cloud computing, as well as for analytics, is of great significance to energy development as well as conservation due to decreased costs as well as customer satisfaction [30]. In addition, IoT devices usually produce a large amount of information that requires an effective storage mechanism for further analysis in order to develop accurate real-time decisions. To deal with this huge data, deep learning assists by providing high precision [31]. Therefore, deep learning, IoT as well as big data analytics are all essential in creating a technologically advanced society.

19 Conclusion

A significant transformation is evident from our daily lives whereby the application of IoT devices as well as technology has significantly increased. Some of the major applications of IoT are smart homes, smart transportation, smart city, and control

of pollution. Moreover, the current advancements in IoT have made researchers, as well as developers, collaborate in improving technology in IoT on a large scale, which will help society to the highest level. Furthermore, in order to achieve these improvements, we must consider various issues and challenges facing IoT, such as environmental issues. In addition, since IoT does not only provide services but also is involved in generating a large amount of information, it is essential to consider big data analytics since it helps in providing precise real-time decisions that could be applied in the establishment of an advanced IoT system.

Further studies can be conducted in multiple areas. For example, IoT functions can be combined to latest technologies such as block chain in healthcare development [32]. An application development study can be conducted through adopting the design science research (e.g., [33–35]). Example application with IoT can also extend the associative knowledge of Big Data Analytics solution design (e.g., in enhancing data sensing and collection practice) in tourism [36–38], education [39], and other associated problem domain.

References

1. Ranger S (2020) What is the IoT? Everything you need to know about the internet of things right now
2. Kumar S, Tiwari P, Zymbler M (2019) Internet of things is a revolutionary approach for future technology enhancement: a review. J Big Data 6(1). https://doi.org/10.1186/s40537-019-0268-2
3. Gatsis K, Pappas G (2017) Wireless control for the IoT. In: Proceedings of the second international conference on internet-of-things design and implementation. https://doi.org/10.1145/3054977.3057313
4. Zhou J, Cao Z, Dong X, Vasilakos A (2017) Security and privacy for cloud-based IoT: challenges. IEEE Commun Mag 55(1):26–33. https://doi.org/10.1109/mcom.2017.1600363cm
5. Minoli D, Sohraby K, Kouns J (2017) IoT security (IoTSec) considerations, requirements, and architectures. In: 2017 14th IEEE annual consumer communications & networking conference (CCNC). https://doi.org/10.1109/ccnc.2017.7983271
6. Behrendt F (2019) Cycling the smart and sustainable city: analyzing EC policy documents on internet of things, mobility and transport, and smart cities. Sustainability 11(3):763. https://doi.org/10.3390/su11030763
7. Jahnke A (2020) The 4 stages of IoT architecture
8. Causevic S, Colakovic A, Haskovic A (2018) The model of transport monitoring application based on the internet of things
9. Nicolescu R, Huth M, Radanliev P, De Roure D (2018) Mapping the values of IoT. J Inf Technol 33(4):345–360. https://doi.org/10.1057/s41265-018-0054-1
10. Park E, del Pobil A, Kwon S (2018) The role of internet of things (IoT) in smart cities: technology roadmap-oriented approaches. Sustainability 10(5):1388. https://doi.org/10.3390/su10051388
11. FRACTTAL (2019) The 9 most important applications of the internet of things (IoT)
12. Babovic ZB, Protic V, Milutinovic V (2016) Web performance evaluation for internet of things applications
13. Home Security Systems. (n.d.) HP news—HP study finds alarming vulnerabilities with internet of things (IoT). WWW.hp.com

14. Sisinni E, Saifullah A, Han S, Jennehag U, Gidlund M (2018) Industrial internet of things: challenges, opportunities, and directions. IEEE Trans Industr Inf 14(11):4724–4734. https://doi.org/10.1109/tii.2018.2852491
15. Yan Z, Zhang P, Vasilakos AV (2014) A survey on trust management for internet of things. J Netw Comput Appl 42:120–134. https://doi.org/10.1016/j.jnca.2014.01.014
16. Pei M, Cook N, Yoo M, Atyeo A, Tschofenig H (2016) The open trust protocol (OTrP). IETF
17. Roman R, Najera P, Lopez J (2011) Securing the internet of things. Computer 44(9):51–58. https://doi.org/10.1109/mc.2011.291
18. Interoperability in Digital Health (n.d.) Reference material|Publications. Publications.iadb.org. https://publications.iadb.org/publications/english/document/Interoperability_in_Digital_Health_Reference_Material_en.pdf
19. Noura M, Atiquazzaman M, Gaedke M (2019) Interoperability in internet of things taxonomies and open challenges. Mob Netw Appl 24:796–809 (References—Scientific Research Publishing). WWW.scirp.org
20. Tzafestas SG (2018) Ethics and law in the internet of things world. Smart Cities 1(1):98–120
21. Pereira C, Aguiar A (2014) Towards efficient mobile M2M communications: survey and open challenges. Sensors 14(10):19582–19608. https://doi.org/10.3390/s141019582
22. Nacera T, Abdelghani C, Karim D, Ahmed-Nacer M (2018) A distributed agent-based approach for optimal QoS selection in web of object choreography
23. Huo L, Wang Z (2016) Service composition instantiation based on cross-modified artificial Bee Colony algorithm. China Commun 13(10):233–244. https://doi.org/10.1109/cc.2016.7733047
24. White G, Nallur V, Clarke S (2017) Quality of service approaches in IoT: a systematic mapping. J Syst Softw 132:186–203. https://doi.org/10.1016/j.jss.2017.05.125
25. Oasis (2012) Web services quality factors Version 1.0 Candidate OASIS Standard 01
26. Saggi MK, Jain S (2018) A survey towards an integration of big data analytics to big insights for value-creation. Inf Process Manage 54(5):758–790. https://doi.org/10.1016/j.ipm.2018.01.010
27. Lee CKM, Yeung CL, Cheng M (2015) Research on IoT based cyber-physical system for industrial big data analytics. In: 2015 IEEE international conference on industrial engineering and engineering management (IEEM). https://doi.org/10.1109/IEEM.2015.7385969
28. Vuppalapati C, Ilapakurti A, Kedari S (2016) The role of big data in creating sense EHR, an integrated approach to create next-generation mobile sensor and wearable data-driven electronic health record (EHR). In: The 2016 IEEE second international conference on big data computing service and applications (BigDataService). https://doi.org/10.1109/BigDataService.2016.18
29. Mourtzis D, Vlachou E, Milas N (2016) Industrial big data as a result of IoT adoption in manufacturing. Proc CIRP 55:290–295
30. Ramamurthy A, Jain P (2017) The internet of things in the power sector opportunities in Asia and the Pacific ADB sustainable development working paper series no. 48. https://www.adb.org/sites/default/files/publication/350011/sdwp-48.pdf
31. Mohammadi M, Al-Fuqaha A, Sorour S, Guizani M (2018) Deep learning for IoT big data and streaming analytics: a survey. IEEE Commun Surv Tutorials 20(4):2923–2960
32. Prokofieva M, Miah SJ (2019) Blockchain in healthcare. Australas J Inf Syst 23:1–22
33. Genemo H, Miah SJ, McAndrew A (2015) A design science research methodology for developing a computer-aided assessment approach using method marking concept. Educ Inf Technol 22:1769–1784
34. Miah SJ, Gammack JG, McKay J (2019) A metadesign theory for tailorable decision support. J Assoc Inf Syst 20(5):570–603
35. Miah SJ, McKay J (2016) A new conceptualisation of design science research for DSS development. In: Proceedings of the 20th Pacific Asia conference on information systems (PACIS 2016), Taiwan. http://www.pacis2016.org/Abstract/ALL/758.pdf
36. Miah SJ, Vu HQ, Gammack J (2019) A big-data analytics method for capturing visitor activities and flows: the case of an Island Country. Inf Technol Manage 20(4):203–221
37. Miah SJ, Vu HQ, Gammack J (2018) A location analytics method for the utilization of geo-tagged photos in travel marketing decision-making. J Inf Knowl Manage 18(1):1–13

38. Miah SJ, Hasan N, Gammack JG (2020) A methodological requirement for designing healthcare analytics solution: a literature analysis. Health Inf J 26(4):2300–2314
39. Miah SJ, Miah M, Shen J (2020) Learning management systems and big data technologies for higher education. Educ Info Technol 25:725–730

Review and Evaluation of Trending SSVEP-Based BCI Extraction and Classification Methods

Bayar Shahab

Abstract The rapid development of technology that has involved neurosciences and human–computer interaction has provided solutions to several problems. Brain-computer interface so-called BCI has opened the door to several new research areas and has given way out to critical issues. It has provided solutions to support paralyzed patients to interact with the outside world. This review work presents the state-of-the-art methods and techniques of feature extraction and classifications. These are the methods used to extract and classify the EEG signals. In another way, the features of interest that we are looking for in the EEG-BCI analyzes. Each of the methods from oldest to newest has been discussed while comparing their advantages and disadvantages. This would create a great context and help researchers understand state-of-the-art methods available in this field, their pros and cons, their mathematical representations, and usage. This work makes a vital contribution to the existing field of study. It differs from other similar recently published works by providing the following: (1) categorization/classifications of the SSVEP-based BCI extraction and classification methods, (2) stating most of the prominent methods used in this field in a hierarchical way, and (3) explaining pros and cons of each method and their performance.

Keywords BCI · SSVEP · SSVEP-based BCI · SSVEP-based extraction · SSVEP-based classification

1 Introduction

Neurotechnology seems to be one of the most dominant fields in this era as technology evolves quicker than ever. One of the most interesting technological improvements is neurotechnology, the science and study of the interaction between the human brain and computers. Initially, brain signals were recorded and analyzed on a printed paper by physicians or neurologists to analyze the brain activities for different purposes

B. Shahab (✉)
University of Rome Tor Vergata, 00133 Rome, Italy
e-mail: bayar.azeez@students.uniroma2.eu

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_6

such as coma and tumors. Nowadays, current and potential neurotechnology applications can support paralyzed patients interacting with the outside world. Considering that various challenges face this field of study, from understanding the topics from the actual brain to the applications that might have safety and security concerns. Understanding the context of this research requires preliminary descriptions and explanations of the methods and techniques used to solve such problems and their evaluations, issues, advantages, and theories behind them. There are several EEG-BCI paradigms as SSVEP, MI, and P300. In this work, the literature and works related to SSVEP have been reviewed and analyzed. It has been discussed how SSVEP can over race other methods due to its least training time required and the higher accuracy for recognition and classification [1]. In the following section, methods related to extraction and classification have been reviewed hierarchically.

2 SSVEP-Based BCI Extraction Methods

SSVEPs are brain oscillations produced by changes in the visual part of the brain with frequencies more than 6 Hz. SSVEP signals carry a variety of features, and it is vital to understand them for making an accurate feature extraction or applying a specific classifier on the signals. It is essential to understand which of the features we are dealing with, their properties, what to consider, and how to deal with the issues that come with them. Features are only a data representation that is embedded inside of the EEG signals. EEG signals generally contain time, frequency/spectral, and space information [2]. The most used feature for SSVEP extraction is the spectral feature. This is done by analyzing the amplitude or power spectrum in the frequency domain. Furthermore, we need to understand the properties of the EEG signals, time of events and activities, need for training or not, and if the features have been analyzed or extracted from several or one channel. Apart from the feature and their properties, some other technical aspects play a significant role in SSVEP extraction; the number of stimulus frequencies, number and choice of channels, extraction method itself, and SNR. In the below subsections, we review most of the commonly used methods for SSVEP extraction while stating their pros and cons, and improvements.

2.1 Independent Component Analyzes—ICA

ICA is a statistical method that focuses on the statistical composition and independence of the signals. It separates the signal sources into independent subcomponents. The idea is that signals from different independent sources, i.e., brain signals, contain EEG data, time, and noise. ICA as a new concept and terminology was firstly introduced by Comon [3]. The paper states that ICA could extend the principal component analysis (PCA) method due to its similarities. Some conditions must be taken into consideration when using the ICA method. The conditions are, firstly, the source

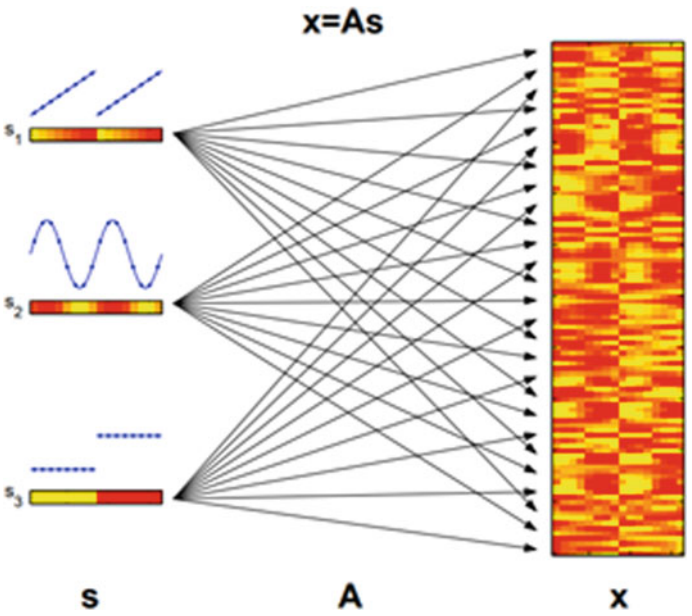


Fig. 1 Mixing function of ICA

signals/components must be statistically independent. Second, the source components must have non-Gaussian distribution. Lastly, the unknown matrix is to be square. Figure 1 shows the mixture of the signals and the components before and after the separation. The output signal X can be multiplied with the inverse of mixture matrix A to recover the original signal. Both these functions must be close to each other to validate the reliability of the applied method, which is the building block of the ICA method.

ICA has been used in several applications [4] to analyze medical images. One of the significant advantages and usage of ICA is to remove noise as in the works of [5, 6]. However, as [3] mentioned, the algorithm can perform the best only if the source signals are non-Gaussian. The author of [7] states that the number of electrodes should be greater than the number of source signals to have the best performance.

2.2 Fast Fourier Transform—FFT

One of the traditional methods for extracting EEG signals is to use the signal's amplitude in the frequency spectrum. FFT transforms the signal from the time domain into the frequency domain to analyze the signal's power spectrum (amplitude). Figure 2 shows a particular subject that is looking at a 15 Hz flickering light. Analyzing the signal in the frequency domain can tell us the subject is looking at that flickering

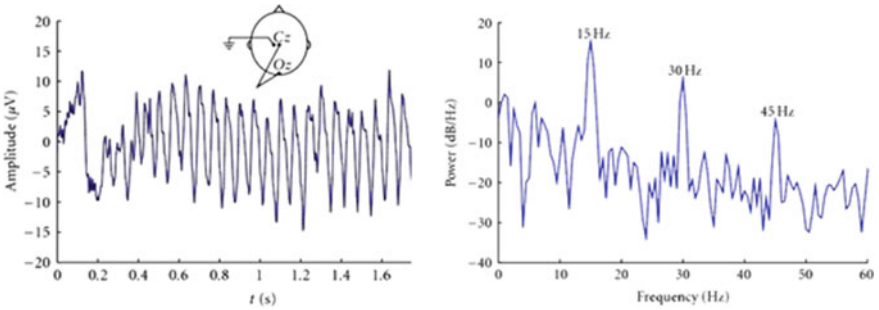


Fig. 2 Signal representation in both time and frequency domains

light by analyzing the amplitude. Hence, we would have an easy signal extraction. FFT, an optimized version of the discrete Fourier transform (DFT), is represented in Eq. 1.

$$F(f) = \sum_{t=0}^{N-1} s(t) e^{\frac{-2i\pi ft}{N}} \tag{1}$$

$s(t)$: is the signal, $t = 1, \ldots, N$ (samples).

FFT has a few advantages; the speed of extraction, and it is easy to use and analyzes for signals. On the other side, it can only perform well when the signals are stationary. Furthermore, as stated by [8], it is sensitive to noise.

2.3 Wavelets Transform—WT

WT is a flexible and powerful signal extraction method used to extract complicated, unstable, and non-stationary signals. In the case of EEG and other signal types like speech with varying frequencies, WT can work better than FFT [9]. WT decomposes the signal into infinite series of smaller parts of so-called wavelets. WT is a time–frequency domain method in which both time and frequency information of the signal are used to extract the information. It uses a flexible and varying window size to adjust to the variances of the signal. The short window size can obtain high-frequency data and long ones for low-frequency data. There are two primary WT categories; continuous wavelet transforms (CWT) for analog and continuous time–frequency analysis and discrete wavelet transform (DWT) for discrete signal analyzes. Another essential feature of the WT is the wavelet type, and it directly affects the performance of the method. Common wavelet types, also called mother wavelets, are Morlet [10], Mexican hat wavelets, and Daubechies. The most common category used in CWT is only signal convolution with wavelets, as shown in Eq. 2.

$$W(s,t) = \int_{-\infty}^{\infty} x(t)\psi*(t)\mathrm{d}t \tag{2}$$

$\psi*(t)$ = wavelet function, $x(t)$ = signal, $W(s,t)$ = wavelet coefficient.

WT is used for extracting non-stationary signals. It has a varying window size that can be used for obtaining different frequency band information. On the other side, the choice of the mother wavelet type affects the performance of the method [8].

2.4 Canonical Correlation Analyzes—CCA

CCA is a multivariate statistical method that recognizes the SSVEP-based signals between two sets of variables that are linear and correlated to each other. The first set of the variables is the EEG signal inputs, and the second is reference sine–cosine waves/signals generated which have the same length as the EEG data. In 2007, it has been used as an SSVEP feature extraction method by [11]. The original CCA method tries to find the linear combination between the coefficients of the two datasets that have the maximum correlation. Let us consider that we have two datasets $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$, with a linear correlation and the same length. X is the EEG signal, and Y is the sine–cosine reference signal. CCA tries to find the linear correlations between $X = \mathrm{XT}\ W_x$ and $Y = \mathrm{YT}\ W_y$ by maximizing the correlation between two datasets. The reference signal is represented as in Eq. 3.

$$Y = \begin{cases} \sin(2\pi f_k t) \\ \cos(2\pi f_k t) \\ \quad\cdot \\ \quad\cdot \\ \sin(2\pi N_h f_k t) \\ \cos(2\pi N_n f_k t) \end{cases}, \quad t = \frac{1}{F_s}, \frac{2}{F_s}, \ldots, \frac{T}{F_s} \tag{3}$$

f_k: stimulus frequency, T: number of sampling points, F_s : is the sampling rate, and N_h is the number of harmonics variables used. CCA will find the weight values W_x and W_y by maximizing the correlation between X and Y variables as in Eq. 4.

$$\max_{W_x W_y} \rho(x,y) = \frac{E\left[x^T y\right]}{\sqrt{E\left[x^T x\right]E\left[y^T y\right]}} \tag{4}$$

In that step, we will have several correlations ρ_i values, from 1 … K values, and the highest correlation would be the maximum classified value. Several improvements to the original CCA method have been done. Most of the methods focus on improving or changing the reference signal of the CCA, such as ITCCA, PCCA,

MwayCCA, L1-MCCA, MsetCCA, CACC, and FBCCA. The original CCA method has the advantage of speed, which is suitable for online computation and has a high accuracy of classifications. However, as stated by [12], CCA sometimes fails to identify the idle and detection stages due to a lack of information while doing the analyzes.

2.5 Hilbert-Huang Transform—HHT

HHT method proposed by [13] to analyze and extract time–frequency information from non-stationary and nonlinear signals. HHT represents any complex dataset into a finite number of intrinsic mode functions (IMFs) using a method called empirical mode decomposition (EMD) by a shifting process iteratively. So, the first stage is the preprocessing using the EMD to decompose the signals into IMFs. As explained by [13], an "IMF is an analytical, self-constructed, well-defined, data-driven function whose amplitudes and frequencies vary with time." HHT is applied to the IMFs to compute the Hilbert spectrum, and the energy in each of the frequency bands can be used as features for SSVEP detection. The apparent advantage of the HHT method is it can classify non-stationary and nonlinear signals. Additionally, it can represent the results in the time–frequency domain. However, [14] explains the mathematical issues of HHT such as the HHT approximation problem, the algorithm to choose the best IMFs, and the adaptive data analyzes method.

2.6 Power Spectral Density Analysis—PSDA

PSDA is one of the traditional methods for SSVEP recognition. It is based on EEG signal analyzes in the frequency domain. The signal is transformed into the frequency domain using Fourier transform, then analyzing the power densities around the stimulus frequencies. PSDA selects the optimum bipolar electrode with the highest SNR value for the K stimulus of targets. The largest frequency with SNR would be computed as in Eq. 5 and acts as a high-pass filter to recognize the features.

$$\mathrm{SNR10}\log_{10} = \left(\frac{nP(f_k)}{\sum_{m=1}^{n/2}(f_k + mf_{\mathrm{res}}) + P(f_k - mf_{\mathrm{res}})} \right) \tag{5}$$

where n: number of frequency points, $P(f_k)$: power density of the stimulus frequencies, mf_{res}: resolution frequency, $(f_k + mf_{\mathrm{res}})$and $P(f_k - mf_{\mathrm{res}})$: power densities around target frequency.

Zerafa et al. [15] states that a single bipolar may affect the performance of PSDA. The time window for analyzing the PSDA method in the frequency domain must be a minimum of 3, limiting the performance. However, it is simple to use, improves

accuracy, and has inter-subject variability of parameters such as location and place of electrodes.

2.7 Least Absolute Shrinkage and Selection Operator (LASSO)

As introduced by [16], LASSO is an estimation method for linear models. The method proposed to solve the issues of the ordinary least squares (OLS) estimation. The issue is estimation accuracy as the coefficient has a significant variance and LASSO shrinks or lowers some coefficients to zero. The interpretation problem has a large number of predictors and subsets of data. LASSO provides an analytical solution with low variance and high interpretations. In 2012, [17] proposed the usage of LASSO for SSVEP recognition. The authors have utilized the advantages of LASSO to extract SSVEP-based signals using the linear regression between the EEG data and square wave signals with different frequencies. Let y be a linear regression model $y \in \mathbb{R}^n$

$$y = X\beta + \varepsilon \tag{6}$$

y: $n \times 1$ vector, $X = (X_1, X_2, \ldots, X_p)$, a $n \times p$ design matrix, ε: a noise vector with the zero mean and constant variance. LASSO estimator would be as in Eq. 7.

$$\hat{\beta} = \arg_{\beta} \min\left(\|y - X\beta\|_2^2 + \lambda\|\beta\|_1\right) \tag{7}$$

where $\|\cdot\|_1$, $\|\cdot\|_2$ denote the l1-norm and l2-norm, λ: is a penalty parameter.

The optimization problem in Eq. 7 can be solved using quadratic programming to define the target frequency. The most substantial contribution of the stimulus frequency would be defined as the target frequency. Zerafa et al. [15] have pointed out some issues regarding the method, such as the ability to analyze multi-channel EEG signals, and each channel needs to be procced separately. The penalty parameter significantly affects the method's performance, and choosing the best parameter can be tricky important to the method. Additionally, using a larger region of the signal can introduce noise, and hence, the SNR for LASSO will decrease. However, it works very well for short data lengths and has low-variance estimates and high interpretation for regression.

2.8 Multivariate Synchronization Index (MSI)

MSI method was proposed in 2014 by [18] for SSVEP recognition. This method estimates the synchronization between the EEG signals and the generated sine–cosine reference signals to find a synchronization index used for recognition. The S-estimator has been used as the index to find the highest stimulus frequency. The largest synchronization index represents the highest correlation between the two sets of data. Hence, the largest reference signal would be marked as the target frequency. Let X be the EEG signals with $N * M$ size and the artificially generated reference signal Y with $2N_h * M$. N is the number of channels, M is the number of samples, N_h is the number of harmonics for Y. The correlation matrix between X and Y is calculated in Eq. 8.

$$C = \begin{bmatrix} C11 & C12 \\ C21 & C22 \end{bmatrix} \tag{8}$$

$C11 = \frac{1}{M} XX^T, C22 = \frac{1}{M} YY^T, C12 = C21 = \frac{1}{M} XY^T$, above equation contains the autocorrelation, which affects the synchronization. To remove the autocorrelation, Eq. 9 represents the linear transform for the same.

$$C = \begin{bmatrix} C11^{-(1/2)} & 0 \\ 0 & C22^{-(1/2)} \end{bmatrix} \tag{9}$$

Equation 9, which is transformed, is then represented in Eq. 10.

$$R = \mathrm{UCU}^T \begin{bmatrix} I_{N*N} & C11^{-(1/2)} C12 C22^{-(1/2)} \\ C22^{-(1/2)} C21 C11^{-(1/2)} & I_{2N_h * 2N_h} \end{bmatrix} \tag{10}$$

Let us consider that, $\lambda_1, \lambda_2, \ldots, \lambda_p$, the eigenvalues of the matrix R, then the normalized eigenvalues can be represented as in Eq. 11.

$$\lambda_i = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i} = \frac{\lambda_i}{\mathrm{tr}(R)} \tag{11}$$

The synchronization index between the two sets of data is shown in Eq. 12.

$$S = 1 + \frac{\sum_{i=1}^{p} \lambda_i \log \lambda_i}{\log(P)} \tag{12}$$

MSI provides good accuracy as it uses nonlinear combinations between the two sets of data. However, it needs a fixed TW for its analyzes.

2.9 Welch's Method

Welch's method was initially introduced by [19] for power spectral estimations. This technique splits the EEG data into N number of segments; then, it calculates the Fourier transform for each segment to compute the modified periodogram. The power spectral estimation would be computed by averaging the modified periodograms as represented in Eq. 13.

$$S\hat{S}(\omega) = \frac{1}{\mathrm{KNU}} \sum_{K=1}^{K} \left| \sum_{K=1}^{K} W(n) * (n + \mathrm{KD}) e^{(-j\omega n)} \right|^2 \tag{13}$$

K: number of segments, N: length of the segments, D: shifted points, W: window function, U is a constant represented in Eq. 14.

$$\hat{S}(\omega) = \frac{1}{N} \left| \sum_{n=1}^{N} W(n) \right|^2 \tag{14}$$

Welch's method reduces noise and artifacts from the signals. Another advantage of the method is that overlapping the signals prevents information loss during the windowing process. However, the length and choice of the window used can affect the performance and accuracy of the classification.

2.10 Minimum Energy Combination—MEC

MEC has been introduced to solve the multi-channel detection of SSVEP responses problem. It has been used by [20] to solve the spatial filtering methods' multi-channel selection problem. The method does not require training and successfully removes noise from the results. It uses a spatial filter to combine multiple electrodes linearly and improve the SSVEP detection results. It is based on the principal component analysis (PCA) method that identifies the signals with high variance. One drawback of the method is the number of selected electrodes that might affect its performance.

2.11 Benchmark, Advantages, and Disadvantages of the Methods

In this section, the list of extraction methods has been presented in a tabular format with a brief description of the technique, stating their advantage(s) and disadvantage(s). Table 1 lists down all the methods reviewed under this section. This can be

Table 1 SSVEP extraction methods with their advantage(s) and disadvantage(s)

Method	Description	Advantage(s)	Disadvantages(s)
ICA	The statistical method separates the source signal into independent components, then mix and recover them	Removes noise and artifacts from EEG signals	• Signals should be non-Gaussian The number of electrodes should be more than the source signals
FFT	Transforms the signal from the time domain into the frequency domain to analyze the power spectrum(amplitude)	• High speed of extraction • Easy to use • Easy to analyze signals	• Works only with stationary signals • Sensitive to noise
WT	Decomposing the signal into infinite series of smaller parts of so-called wavelets	• Extracting non-stationary signals • Varying window size	Choice of wavelet type affects the performance
CCA	Statistical multivariate method that recognizes the SSVEP-based signals between two sets of variables that are linear and correlated to each other	• Good speed for online computation • High accuracy of calculation	Fails to identify the idle and detection stages due to lack of information while doing the analyzes
HHT	Represent signals into a finite number of (IMFs) using (EMD)	Extracting non-stationary and liner signals	• Approximation problem • Algorithm to choose best IMFs
PSDA	FFT transform on signals and analyzing power densities around stimulus frequencies	• Simple to use and improves accuracy • Inter-subject variability	• Single bipolar • TW in the frequency domain must be at a minimum of 3
LASSO	An estimation/prediction method for linear models. Provides an analytical solution with low variance and high interpretations	• Works well for short data lengths • Low-variance estimates and high-interpretation regression	• Unable to analyze multi-channel EEG signals • Penalty parameter choice • Uses large region of signal
MSI	Estimates the sync between the EEG signals and generated reference signals	Has good accuracy rate	It needs a fixed TW
Welch's method	Splits the EEG data into N number of segments, then it calculates the Fourier transform for each of these segments	• Reduces noise and artifacts from signals • Overlapping of signals	Length and choice of TW can affect the performance

(continued)

Table 1 (continued)

Method	Description	Advantage(s)	Disadvantages(s)
MEC	Solves the multi-channel selection and detection problem that exist in the spatial filtering methods	• It does not require training and calibration • Removes noise	A selected number of electrodes can affect its performance

used as a benchmark for other researchers while analyzing and working with similar subjects. Enormous efforts have been spent on this work to collect and provide an easy benchmark and review for future researchers. Additionally, intense research has been conducted to analyze each of the methods, understand them, break them down into smaller parts, state their mathematical representations, and finally figure out their advantages and disadvantages.

3 SSVEP-Based BCI Classification Methods

Classification is a step usually followed by feature extraction. It can be defined as recognizing the class or group of the features that have been recognized and mapping them to their appropriate feature vector. The role of the classifier is as important as the feature extractors. It can also be defined as the mapping between the input features and their appropriate labels, which are the subject's actual targets or mental state. The classifier's job would be to assign the target group. The classifier can be linear or nonlinear according to the classification problem [21]. In the below subsections, common methods for SSVEP classification have been reviewed, stating their pros and cons.

3.1 Bayesian Classifier

The Bayesian classifier is a statistical classifier that classifies the feature vectors into a group according to their probability and estimation. There would be several feature vectors estimations, and the one with the highest probability will be defined as the feature vector. Bayes is motivated by the original Bayes theorem as in Eq. 15.

$$P_{(y|x)} = \frac{P_{(y)} P_{(y|X)}}{P_{(x)}} \tag{15}$$

The values of the Bayesian classifier are independent variables. Naïve Bayes, a restricted version of the Bayesian classifier, uses nodes instead of independent variables, and the edges of the nodes define the dependence between the variables. As stated by [22], the main advantages are its simplicity of usage and flexibility of

changing the variables to have different estimations and its fast speed. The disadvantage is that both Bayesian and Naïve Bayes expect the attributes to be independent of each other.

3.2 *Liner Discriminant Analyzes—LDA*

The original discriminate analyzes have been presented by Fisher [23]. The idea was to classify an object into one of the two defined groups. Later, multiple or so-called dimensional groups have been solved using the same or improved version of the discriminate analyzes. LDA aims to classify the input feature space into several classes using the hyperplane or decision boundary. This is done by transforming the input data into another space by reducing the number of dimensions in the new space of decision using a threshold value to determine the class number/grouping. Equation 16 represents the hyperplane formula of the transformation.

$$m(x) = W^T x + W_o \tag{16}$$

W: weight vector, x: input feature vector, and W_o: threshold value.

One of the main problems that LDA solves is removing the overlapping of feature points between the different classes. Will have a good separation between the classes using the thresholding value as in Fig. 3. Low computational cost and easy use of the method as well as the accuracy are the pros of the method. However, the disadvantages or limitations of this method are the assumptions that the class representations have Gaussian distribution and equal covariances [24].

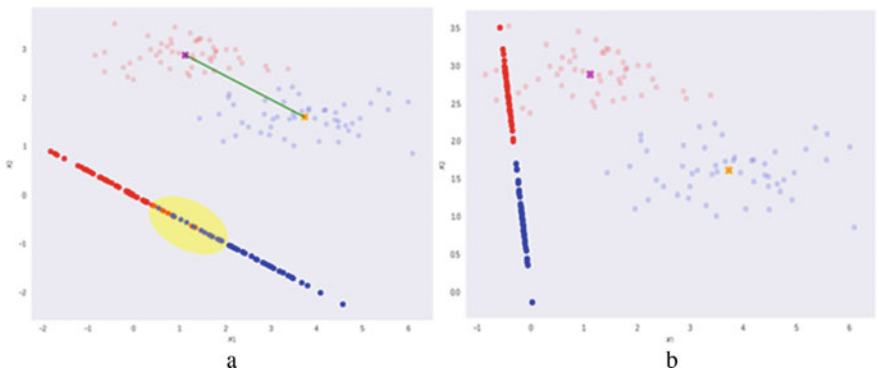


Fig. 3 **a** Overlapping issue after hyperplane projection, **b** LDA's solution of separation [4]

3.3 Support Vector Machine—SVM

SVM is one of the most used classifiers in machine learning and regression. It was introduced first by [25] a group of AT&T. SVM is so-called a maximum margin classifier. In SVM, the feature vector would be classified into several hyperplanes. The job of SVM is to find the optimal hyperplane using the maximum margin between the sets of the data. Maximum margin means to have the farthest distance between the different sets of data. The structure of SVM is based on the kernel function, which can be linear or nonlinear. The kernel function will determine the decision boundary or the hyperplane separation and can be linear, cubic, polynomial, Gaussian. SVM is simple to use and can cope with high-dimensional spaces, and the choice of kernel function can be seen as one of the pros of this method. Nevertheless, the choice of the kernel function can be a crucial and sensitive task.

3.4 K-Nearest Neighbor—KNN

K-nearest neighbor has been used for solving classification and regression issues. It has been used to classify the feature points. The basic concept of this method is finding similar feature points that are near or close to each other. There might be some features flocked in different classes of data; however, by finding its nearby feature points, they might be classified as a separate class of features, which is done by finding its nearest neighbors. The traditional Euclidian distance can calculate the distance between the nearby neighbors as in Eq. 17.

$$\text{distance}(x, y) = \sqrt{\sum_i (x_i - y_i)^2} \tag{17}$$

The most important part of this method is choosing the K-number that identifies the nearby neighbors between the similar features. K-number should always be an odd number to remove any uncertainty in the classification. A traditional method to choose the best K-number is to run the algorithm several times using different values and see which one will have a minimal error rate. Usually, less K-number is less stable and the more K-number is the more stable will be. The KNN method has the advantage of having a minimal error rate for decisions and is relatively easy to implement. On the contrary, it is not so commonly used in the BCI community [26].

3.5 Extreme Learning Machines—ELM

ELM was first introduced by [27] in 2004, inspired by the common multilayer neural network perceptron. The structure of ELM consists of a single input layer, a single

hidden layer, and an output layer. Different methods can determine the number of neurons per layer, and further research has used different algorithms for determining this number, such as the cross-validation method. The ELM method assigns the input weights, the hidden layer, and the bias values based on continuous probability distribution functions. ELM has also been used with CCA to enhance single-label classification and make it multi-label classification. ELM has the advantage of being simple to use, and it is fast to learn. While ELM is a fast method to compute the results, using only a single hidden layer cannot always guarantee accurate and best results.

3.6 Convolutional Neural Network—CNN

In a work done by [28] in 2014, authors have used a CCN to control the motion and movement in a virtual reality environment (VE) and have great accuracy and results. Another significant research has been conducted by [29], which has used CCN for SSVEP classification under an ambulatory environment. The authors have used various CNN networks with different layers, namely CCN-1 and CCN-2, and have outstanding results. In 2018, [30] have used CCN to classify SSVEP signals without any preprocessing and feature extraction techniques. The convolution layer will transform the data into feature points that eliminate preprocessing and signal extraction. Using a specific function in the final layer will classify the input features into their targeted labels. In 2019, [31] authors have used the same approach to identify asynchronous SSVEP signals. The main and obvious advantage of this method is the accuracy calculation which is significantly higher than other methods. Another point is the elimination of some preprocessing and extraction sometimes. Designing the network and deciding on the number of layers can be a difficult task and need careful attention that affects the accuracy calculation.

3.7 Benchmark, Advantages, and Disadvantages of the Methods

Same as Sect. 2.11, in this section, the list of classification methods has been presented in a tabular format with a brief description of the technique, stating their advantage(s) and disadvantage(s). Table 2 lists down all the methods reviewed under this section.

Table 2 SSVEP-based classification methods with their advantage(s) and disadvantage(s)

Method	Description	Advantage(s)	Disadvantages(s)
Bayesian classifier	A statistical classifier that classifies the feature vectors into a group or class according to their probability and estimation	• Easy to use • The flexibility of changing variables • Has fast speed	The method expects the attributes to be independent of each other
LDA	classify an object into multiple defined groups	• Easy to use • Low-computational cost • High accuracy	Class representations should have Gaussian distributions and equal covariances
SVM	Feature vector would be classified into several hyperplanes, and the job of SVM is to find the optimal hyperplane using the maximum margin between the sets of the data	• Simple to use • Can cope with high-dimensional spaces • Choice of kernel	Choice of the kernel function can be a crucial and sensitive task
KNN	Finding similar feature points that are near or close to each other	• Ease to implement • Minimal error rate	It is not so commonly used in the BCI community
ELM	Assigns the weights of the input, the hidden layer, and the bias values based on some continuous probability distribution functions	• Simple to use • Fast to learn	• The choice of using only a single hidden layer cannot always guarantee accurate and best results
CNN	Using deep neural networking approach to solve the classification problem	• High accuracy • Removing some of the preprocessing methods	• The complexity of designing the network • Deciding the number of layers

4 Conclusion and Discussion

EEG-based BCI field has mapped the road for special applications and has created opportunities and hopes for humans to achieve things that were difficult to be attained. With the power of the brain integrated or interacting with fast, modern. Robust computer systems, humans now can solve various issues such as helping paralyzed patients to interact with the world, memory rehabilitation of patients, and perhaps shortly humans with chips integrated into their brain to be more intelligent, more potent, and help them with brain issues. This paper gives a comprehensive review and analysis of the SSVEP-based BCI extraction and classification techniques. Additionally, it analyzes and evaluates the pros and cons of the methods that have been

discussed. Finally, gaps that exist in each of the methods have been specified. The paper mentioned how EEG-based BCI methods have improved with the technology enhancements that led to the founding and improvements of several EEG feature extraction methods as CCA, ICA, FFT, WT, and HHT.

On the other hand, feature classification methods have seen several improvements, and lots of methods do exist, including Bayes classifier, LDA, SVM, KNN, ELM, and CNN. Each of these methods serves and solves certain types of problems in the BCI community and field. Tables 1 and 2 show an overall taxonomy of the extraction and classification techniques in a tabular format with a brief description and stating their pros and cons. Thus, any researcher could use this review as a benchmark for making decisions about each method. This work services future researchers to have a strong knowledge of the previous related works. It shows, analyzes, and evaluates existing and state-of-the-art methods in the related area of the research to avoid duplication and repetition.

References

1. Zhang W, Tan C, Sun F, Wu H, Zhang B (2018). A review of EEG-based brain-computer interface systems design. Brain Sci Adv 156–167
2. Vialatte F, Maurice M, Dauwels J, Cichocki A (2010) Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. Prog Neurobiol 418–438
3. Pierre Comon (1994) Independent component analysis, a new concept? Signal processing, vol 36. Elsevier, pp 287–314. Ff10.1016/0165-1684(94)90029-9ff. Ffhal-00417283f
4. Isomura T, Toyoizumi T (2016) A local learning rule ICA. Nature
5. Gramfort A, Luessi M, Larson E, Engemann D, Strohmeier D, Brodbeck C, Hämäläinen M (2013) MNE software for processing MEG and EEG data. Neuroimage 15
6. Mahajan R, Morshed B (2015) Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis and wavelet-ICA. IEEE J Biomed Health Inform 158–165
7. Naik G, Kumar D (2009) Determining number of independent sources in undercomplete mixture. EURASIP J Adv Sig Process 5
8. Al-Fahoum A, Al-Fraihat A (2014) Methods of EEG signal features extraction using linear. ISRN Neurosci 7
9. Zhang Z, Li X, Deng Z (2010) A CWT-based SSVEP classification method for brain-computer. In: International conference on intelligent control and information processing, Dalian, p 6
10. Grossmann A, Morlet J (1984) Decomposition of hardy functions into. SIAM J Math Anal 723–736
11. Lin Z, Zhang C, Wu W, Gao X (2007) Frequency recognition based on canonical. IEEE Trans Biomed Eng 1172–1176
12. Poryzala P, Materka A (2014) Cluster analysis of CCA coefficients for robust detection of the asynchronous SSVEPs in BCI. Biomed Sig Process Control 201–208
13. Huang N, Shen Z, Long S, Wu M, Shih H, Zheng Q, Liu T et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: Royal society. The Royal Society Publishing, London, pp 903–995
14. Huang NE (2005) Introduction to the Hilbert–Huang transform and its related mathematical problems. Interdisciplinary Math Sci 1–26.https://doi.org/10.1142/9789812703347_0001
15. Zerafa R, Camilleri T, Falzon O, Camilleri K (2018) To train or not to train? A survey on training of feature extraction method for SSVEP-based BCIs. J Neural Eng 29

16. Tibshirani R (1996) Regression shrinkage and selection via the lasso. J Royal Stat Soc Ser B (Methodological) 267–288
17. Zhang Y, Jin J, Qing X, Wang B, Wang X (2012) LASSO based stimulus frequency recognition model for SSVEP BCIs. Biomed Sig Process Control 104–111
18. Zhang Y, Xu P, Cheng K, Yao D (2014) Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface. J Neurosci Methods 32–40
19. Welch P (1967) The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE
20. Friman O, Volosyak I, Gräser A (2007) Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. IEEE Trans Biomed Eng 742–750
21. Carvalho S, Costa T, Uribe L, Soriano D, Yared G, Coradine L, Attux R (2015) Comparative analysis of strategies for feature extraction and classification in SSVEP BCIs. Biomed Sig Process Control 34–42
22. Soltani H, Einalou Z, Maghooli K (n.d.) Classification of SSVEP-based BCIs using genetic algorithm. Research Square, Tehran
23. Huberty C (1975) Discriminant analysis. Rev Educ Res 543–598
24. Bishop C (2006) An illustrative introduction to Fisher's linear discriminant. In Bishop C (ed) Pattern recognition and machine learning (information science and statistics). Springer, London, p 738
25. Yayik A, Kutlu Y (2017) Online LDA based brain-computer interface system to aid disabled people. Nat Eng Sci 13
26. Boser B, Guyon I, Vapnik V (1992) A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on computational learning theory, Pittsburgh
27. Singla R (2018) SSVEP-based BCIs. In: Larrivee D (ed) Evolving BCI therapy—engaging brain state dynamics. Intechopen. https://doi.org/10.5772/Intechopen.75693. Available from https://www.intechopen.com/chapters/61167
28. Huang G-B, Zhu Q-Y, Siew C-K (2014) Extreme learning machine: a new learning scheme of feedforward neural network. In: International joint conference on neural networks (IJCNN). IEEE Xplore, Budapest, pp 985–990
29. Bevilacqua V et al (2014) A novel BCI-SSVEP based approach for control of walking in virtual environment using a convolutional neural network. In: 2014 international joint conference on neural networks (IJCNN), pp 4121–4128.https://doi.org/10.1109/IJCNN.2014.6889955
30. Kwak N-S, Müller K-R, Lee S-W (2017) A convolutional neural network for steady-state visual evoked potential classification under ambulatory environment. PLOS One 12(2):E0172578.https://doi.org/10.1371/Journal.Pone.0172578
31. Aznan N, Bonner S, Connolly J, Moubayed N, Breckon T (2018) On the classification of SSVEP-based dry-EEG signals via convolutional neural networks. In: 2018 IEEE international conference on systems, man, and cybernetics (SMC), pp 3726–3731.https://doi.org/10.1109/SMC.2018.00631
32. AVSS 2007 (2007) In: 2007 IEEE international conference on advanced video and signal based surveillance. Retrieved from eecs.qmul: http://www.eecs.qmul.ac.uk/~andrea/avss2007_d.html
33. Vacavant A, Chateua T, Wilhelm A, Lequievr L (2012) A benchmark dataset for outdoor foreground/background extraction. In: ACCV 2012, workshop: background models challenge, Korea
34. Weisstein E (n.d.) Distance. Retrieved from MathWorld—a wolfram web resource. http://mathworld.wolfram.com/Distance.html
35. Math is Fun (n.d.) Percentage change. Retrieved from https://www.mathsisfun.com/numbers/percentage-change.html

Finite Element of Biomechanical Model of the Human Myocardium from a Cardiac MRI Images

Awadi Rania, Narjes Benameur, Tesnim Kraiem, and Salam Labidi

Abstract Biomechanical models of the myocardium provide more details of the heart behavior and several biomechanical parameters. Thus, biomechanical heart models are important for improving clinical treatment and interventions for patients with heart failure. The aims of this study are to present a biomechanical human left ventricle (LV) models that are derived from clinical imaging data of 20 healthy subjects. End-systolic volume (ESV), end-diastolic volume (EDV), and end-diastole wall thickness from 20 health subjects were computed using cardiac CMR data and personalized cardiac modeling. The results reveal that the computed parameters are in accordance with the normal values of healthy subjects. The outcome of this study suggests that the proposed 3D model of the LV is able to describe the physiological function of the heart and to differentiate between normal and pathological heart function.

Keywords Left ventricle model · Cardiac magnetic resonance imaging · Finite element · Biomechanical · Simulation · Medical imaging · Modeling

1 Introduction

Myocardial infarction is a heart disease that occurs when the heart does not obtain enough blood and oxygen. This dysfunction usually leads to severe myocardial damage [1]. Post-myocardium infarction alters the mechanical parameters of the heart. Unfortunately, the radiologists are not able to detect these mechanical changes. In this context, computational mechanical models of LV is a useful tools in diagnosis and treatments for heart failure. The estimation of myocardium biomechanical

A. Rania (✉) · N. Benameur · S. Labidi
Research Laboratory of Biophysics and Medical Technologies, Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, LR13ES07, 1006 Tunis, Tunisia
e-mail: rania.awadi@istmt.utm.tn

T. Kraiem
Laboratory of Applied Mechanics and Engineering, National Engineers School of Tunis, University of Tunis El Manar, LR-11-ES19, 1002 Tunis, Tunisia

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_7

parameters and properties are based on the myocardium geometry, loading conditions and myocardium materials [2]. Cardiac magnetic resonance (CMR) imaging provides some information that can be used to model biomechanical models like myocardium geometry and material properties [3].

In recent years, several efforts have been developed to compute mechanical models of the myocardium [4–8]. These models allow the estimation of biomechanical parameters such as contractility, stiffness and deformation. These parameters depend on two behavior: active and passive behavior.

The active behavior is created by the sarcomere, the basic contractile unit in myocytes, at its resting length; it is the required minimum contractile function to meet the body's blood demand. While passive behavior is the correlation among myocardium stress and myocardium strain. It presents the hyperelastic properties of the myocardium and is a passive element of diastolic function [9]. Table 1 summaries different biomechanical models (LV geometry, passive and active laws, and loading condition).

2 Methods

The LV biomechanical model allows to identify diastolic and systolic mechanical parameters of the myocardium. Many models can be used to identify these parameters. According to Gao et al. [18], they can be composed of five sections: LV geometry, LV reconstruction (3D model), materials, the finite element meshing, and boundary limits.

2.1 Dataset

The training cine-MRI dataset of 20 healthy subjects were employed in this work. We used cine-MRI sequences; each sequence is formed of 25 images representing a cardiac cycle. These data were collected from the Principle Military Hospital of Instruction (HMPIT) of Tunis (Tunisia). All images were derived from machine 3 T (Siemens Medical solution, Erlangen, Germany). The local ethics committee provided the study protocol, and the volunteers give written informed consent.

Table 1 Comparative table that presents different modeling studies

Study	Subject number	Biomechanical modeling		Constitutive law		Boundary
		Reference geometry/modality	Motion: modality/quantification	Passive	Active	
Gao et al. [13]	38	LV at ED/CMR	Cine-CMR/estimate regional circumferential strain	H,O law	(T_{req})	Imposed EDP
Walker et al. [14]	5 sheep	LV at end isovolumic relation/CMR	Tagged CMR/ES strain	F,G law	(T_{max})	EDP/ESP measured through catheterization
Rumindo et al. [15]	21	LV at ES/CMR	Cine-CMR/feature tracking for circumferential strain	Simplified FG law	(T_{max})	Fixed base, boundary-free epicardium
Asner et al. [16]	6	LV at ES/CMR	3D tagged MRI/displacements from image registration toolkit (IRTK)	H,O law	Active tension parameters $\alpha(t)$	Endocardial volumes and basal displacements EDP imposed
Marchesseau et al. [17]	11	LV and RV/CMR	Cine-CMR/incorporating an elastic and incompressible regularize into the registration	Hyperelastic material (Mooney-Rivlin) BSC approach	BSC approach	Endocardial volumes imposed EDP
Wenk et al. [18]	2 sheep	LV/echocardiography, MRI	Echocardiography, MRI/incorporating the leaflets of the mitral valve	F,G law	(T_{max})	Extended basal nodes of the LV

(continued)

Table 1 (continued)

Study	Subject number	Biomechanical modeling		Constitutive law		Boundary
		Reference geometry/modality	Motion: modality/quantification	Passive	Active	
Xi et al. [3]	3	LV at BD/CMR	Cine-CMR/displacements from image registration toolkit (IRTK)	F,G law	Length—depende nt active term	LV pressure measured

End-systole (ES), end-diastole (ED), beginning diastole (BD), end-systolic pressure (ESP), end-diastolic pressure (EDP), end-systolic volume (ESV), end-diastolic volume (EDV), feature tracking (FT), cardiac magnetic resonance (CMR), Fung-Guccione law (F,G) [10], Holzapfel-Ogden law (H,O) [11], and Bestel Clement-Sorine model (BSC) [12]

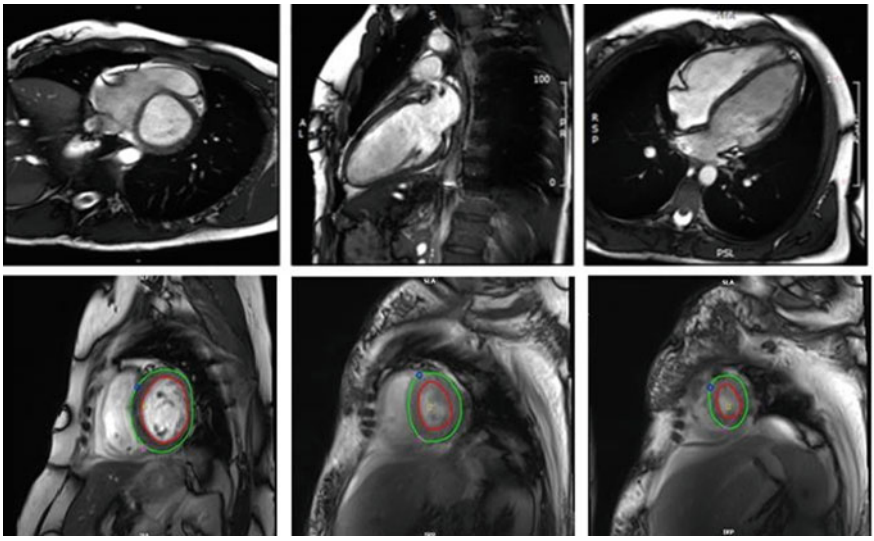


Fig. 1 Top: cardiac MRI images. Bottom: represent segmentation results, respectively. Green: LV epicardium, red: LV endocardium

2.2 Numerical Simulation

2.2.1 Segmentation

Short-axis and the 2-chamber and 4-chamber long axis images were chosen for LV segmentation. We used a Circle Cardiovascular Imaging software (CVI42). The segmentation results are shown in Fig. 1.

2.2.2 3D Surface Reconstruction

Then, we import the segmentation results into a commercial software Ls-Prepost, to assembly LV epicardium and the endocardium (see Fig. 2a).

Assign points of the endocardium and the epicardium in three-dimensional (3D) area are then immediately created from these contours (Fig. 2b). We used HyperWorks software (© Altair 2019) to generate the LV walls reconstruction (Fig. 2c).

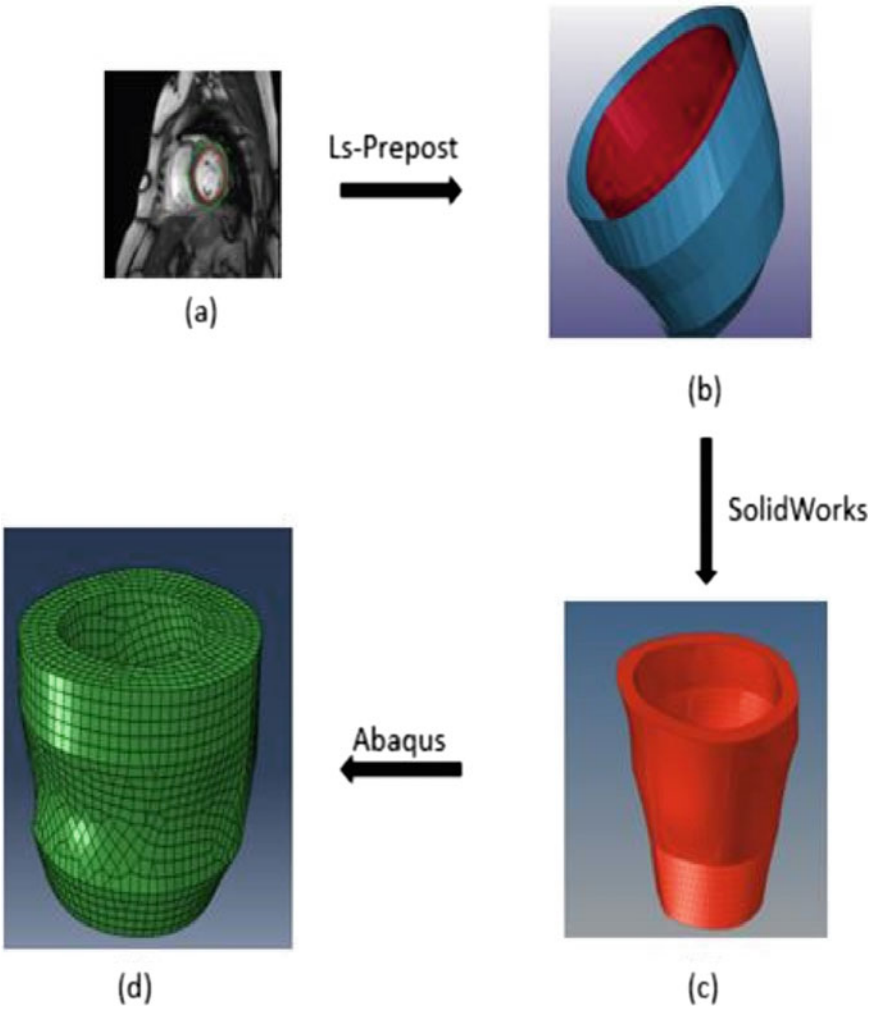


Fig. 2 **a** Digitizing the MR images. **b** Assembly LV epicardium and the endocardium. **c** LV walls reconstruction. **d** 3D volume meshes for LV

2.3 3D Cardiac model

One of the principal functions of LV segmentation is 3D cardiac modeling. Accured to the SolidWorks result, we used Abaqus (©Dassault Systems) to create 3D volume mesh. The hexahedral mesh for each subject composed of 4576 C3D8R elements (Figs. 2d and 3). This work, the ES geometry has been chosen as reference. 3745 ± 450 C3D8R.

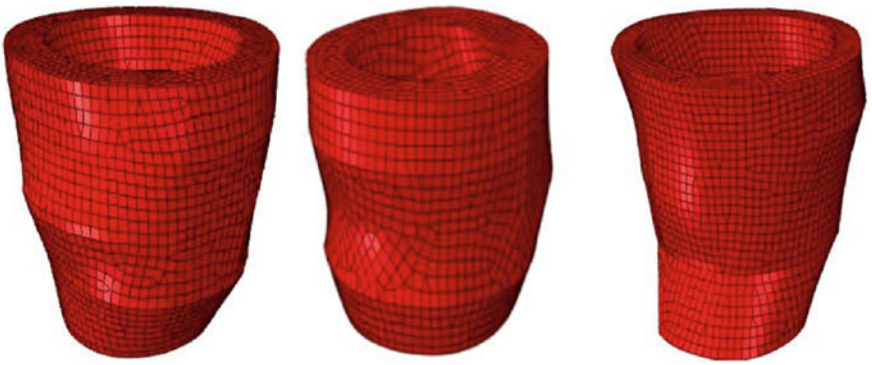


Fig. 3 LV models at ES for three healthy subjects

Table 2 Normal LVED and LVES volumes according to gender

	LVED (ml)	LVES (ml)
Male	62–120	22–66
Female	58–103	26–82

3 Evaluation

To evaluate our work, a clinical metrics of LV simulated from the mechanical models: end-diastolic/end-systolic volume and ED wall thickness. The metrics obtained from the LV simulated is compared to the normal values obtained by the literature. Table 2 displays the normal LVED and LVES normal volumes according to gender [19].

The Fig. 3 shows the results of LV modeling for three healthy subjects. For the 20 cases, we computed LV wall thickness progressively from base to the apex.

4 Results and Discussion

Table 3 presents the results of the LV metrics on the 20 volunteers. For the healthy subject, the normal values range of end-systolic and end-diastolic volumes were determined to be 22–82 ml and 58–120 ml, respectively, for the two gender. We obtained with personalized simulated (subject #8) end-systolic and end-diastolic volumes, respectively, 29 ml and 74 ml; those are in 4% of the normal values. For the LV wall thickness, the mean normal value (subject #8) from base to the apex was 10.3 ± 1.0 mm for base, 10.2 ± 0.9 mm, and 7.3 ± 1.0 mm for apex. Using our personalized simulation, we obtained LV wall thickness of 8.23 mm for the same patient.

Table 3 LV clinical metrics for 20 subjects, 'simulated'

Subject		#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Volume (ml)	End-diastole	59	73	105	82	105	102	64	74	94	84
	End-systole	24	34	56	36	56	64	25	29	40	43
ED wall thickness (mm)		7.81	7.98	6.71	12.3	8.66	9.88	7.8	8.23	7.72	8.80
Subject		#11	#12	#13	#14	#15	#16	#17	#18	#19	#20
Volume (ml)	End-diastole	86	101	75	66	88	92	102	99	101	76
	End-systole	42	72	34	32	42	37	54	32	58	32
ED wall thickness(mm)		9.19	9.80	8.48	8.72	9.81	7.33	10.20	9.91	9.30	8.81

5 Conclusion

In this article, we modeled a biomechanical human left ventricle in healthy subjects. LV end-systolic and end-diastolic volumes and wall thickness from 20 subjects were estimated based on CMR images and biomechanical modeling. The modeling method was largely evaluated by comparing the simulated values to the normal values. We aim in further study to develope a LV model for patients with ischemic myocardium and estimate their biomechanical parameters.

References

1. Richardson WJ et al (2015) Physiological implications of myocardial scar structure. Compr Physiol 5(4):1877–1909
2. Wang VY, Nielsen PM, Nash MP (2015) Image-based predictive modeling of heart mechanics. Annu Rev Biomed Eng 351–383
3. Xi J, Lamata P, Niederer S, Land S, Shi W, Zhuang X et al (2013) The estimation of patient-specific cardiac diastolic functions from clinical measurements. Med Image Anal 17(2):133–146
4. Baillargeon B, Rebelo N, Fox D, Taylor R, Kuhl E (2014) The living heart project: a robust andante grative simulator for human heart function. Eur J Mech A Solids 48:38–47
5. Nash MP, Hunter PJ (2000) Computational mechanics of the heart. J Elast 61:113–141
6. Guccione JM, Moonly SM, Moustakidis P, Costa KD, Moulton MJ, Ratcliffe MB (2001) Mechanism underlying mechanical dysfunction inner the border zone of left ventricular aneurysm: a finite element model study. Ann Thorac Surg 71(2):654–662
7. Wang H, Gao H, Luo X, Berry C, Griffith B, Ogden R, Wang T (2013) Structure-based finite strain modelling of the human left ventricle in diastole. Int J Numer Method Biomed Eng 29:83–103
8. Gao H, Carrick D, Berry C, Griffith B, Luo X (2014) Dynamic finite-strain modelling of the human left ventricle in health and disease using an immersed boundary-finite element method. IMA J Appl Math 79:978–1010
9. Gao H, Aderhold A, Mangion K, Luo X, Husmeier D, Berry C (2014) Advances in computational modelling for personalised medicine after myocardial infarction. J R Soc Interface 14(132):23

10. Gao H, Aderhold A, Mangion K, Luo X, Husmeier D, Berry C (2017) Changes and classification in myocardial contractile function in the left ventricle following acute myocardial infarction. J R Soc Interface 14(132):3–6
11. Guccione JM, McCulloch AD, Waldman LK (1991) Passive material properties of intact ventricular myocardium determined from a cylindrical model. J Biomech Eng 113(1):42–55
12. Holzapfel GA, Ogden RW (2009) Constitutive modelling of passive myocardium: a structurally based framework for material characterization. Philos Trans A Math Phys Eng Sci 367(1902):3445–3475
13. Bestel J, Clément F, Sorine M (2001) A biomechanical model of muscle contraction. In: Niessen WJ, Viergever MA (eds) Medical image computing and computer-assisted intervention MICCAI 2001. Springer, Berlin Heidelberg, pp 1159–1161
14. Walker JC, Ratcliffe MB, Zhang P, Wallace AW, Fata B, Hsu EW et al (2005) MRI-based finite-element analysis of left ventricular aneurysm. Am J Physiol Heart Circ Physiol 289(2):692–700
15. Rumindo G, Ohayon J, Croisille P, Clarysse P (2020) In vivo estimation of normal left ventricular stiffness and contractility based on routine cine MR acquisition. IPEM, 1350–4533
16. Asner L, Hadjicharalambous M, Chabiniok R, Peressutti D, Sammut E, Wong J et al (2017) Patient-specific modeling for left ventricular mechanics using data-driven boundary energies. Comput Methods Appl Mech Eng 314:269–295
17. Marchesseau S, Delingette H, Sermesant M, Cabrera-Lozoya R, Tobon-Gomez C, Moireau P et al (2013) Personalization of a cardiac electromechanical model using reduced order unscented kalman filtering from regional volumes. Med Image Anal 17(7):816–829
18. Wenk JF, Zhang Z, Cheng G, Malhotra D, Bolton GA, Burger M, Suzuki T, Saloner D, Wallace A, Guccione J et al (2010) First finite element model of the left ventricle with mitral valve: insights into ischemic mitral regurgitation. Ann Thorac Surg 89:1546–1554
19. Cain AP, Ahl R, Hedstrom E et al (2009) Age and gender specific normal values of left ventricular mass, volume and function for gradient echo magnetic resonance imaging: a cross sectional study. BMC Med Imaging

Comparative Analysis Between Macro and Micro-Accuracy in Imbalance Dataset for Movie Review Classification

Nur Suhailayani Suhaimi, Zalinda Othman, and Mohd Ridzwan Yaakub

Abstract Classification for multi-class dataset provides exciting and explorative domain to be studied in data science domain. And yet, the challenges of measuring the accuracy of multi-class performance rise an issue worth detailed research to be explored. Due to multi-class accuracy may be lower due to imbalance dataset, this paper aimed to analyze the usage of macro and micro-accuracy in classifying text data with multi-class label. This research focused on text data of movie reviews being classified by three multi-class classifier which are Naïve Bayes (NB), Support Vector Machine (SVM), and Random Forest (RF). We set five performance measure to be analyzed; recall, precision, f-score, sensitivity and specificity with regards of micro and macro-accuracy. We successfully yielded a significant result of comparative analysis where average micro-accuracy (87.3%) produced 14.8% higher than macro-accuracy (72.5%) for imbalance dataset. Result also shown a significant gap between balanced and imbalanced dataset. For further analysis, the flexibility of class label in multi-class may be studied to obtain the changing of learning behavior of the classifier as future work.

Keywords Multi-class classification · Macro and micro-accuracy · Text classification

1 Introduction

Measuring accuracy for multi-class classification toward stream dataset could be very challenging. The accuracy for classification usually used as a metric to indicate performance of a classifier by taking the value of correct classifications using two ways which are independent test set or using variation in cross-validation. However,

N. S. Suhaimi (✉)
Faculty of Computer and Mathematical Sciences, University of MARA Technology, 73000 Merlimau, Melaka, Malaysia
e-mail: suhailayani@tmsk.uitm.edu.my

Z. Othman · M. R. Yaakub
Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, The National University of Malaysia, 43600 UKM Bangi, Selangor, Malaysia

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_8

classification accuracy alone may not sufficient when dealing with multi-class and imbalance dataset. Classification accuracy is developed by practitioners on dataset with an equal class distribution, hence dealing with skewed class distribution raised issues and problems. Therefore, we initiate a study to analyze the influence of other metric such as macro-accuracy, micro-accuracy, recall, precision, f1-score, sensitivity and specificity for multi-class classification with specific text classification datasets. A macro-accuracy computes the metric for each class separately and treat all classes equally, while a micro-accuracy will compute the aggregation of all classes to produce the average metric. Many researchers [8–10] suggest the used of multiple metrics for multi-class due to its equality in calculating the sum of all true positives by each class available. Authors [11, 12] claimed macro and micro-accuracy metrics are the best option along with other metrics such as precision, recall, f1-score, sensitivity and specificity. For some reason, the hybridization of methods, approaches and metrics yield better results [3]. To define the fraction of correct predictions for a particular class, we use precision metric, while to compute the fraction of instances of a class that were correctly predicted, we use recall value. We use average accuracy, which is the value of correctly classified instances in the sum of one-vs-all matrices matrix. In information retrieval, the positive predictive value is called precision, and sensitivity is called recall. The name sensitivity comes from the statistics domain as a measure for the performance of a binary classification, while recall is more related to the Information Engineering domain. Sensitivity refers to the proportion of classified data with positive result on test set out of the set that actually have the condition. Specificity refers to the proportion of classified data as a negative result on this test out of those who do not actually have the condition. As for classifier algorithm, we chose three algorithm which are suitable for multi-class classification [1, 2, 7]; Support Vector Machine (SVM), Naïve Bayes (NB) and Random Forest (RF). Multi-class SVM works well to assign classes to instances by using support vector machines, where the classes are derived from a finite set of several variables. The dominant approach is to minimize the single multi-class classification into multiple binary classification processes. SVMs are helpful in text classification [1], as their application can significantly reduce the need for categorizing training instances in both the standard inductive and transudative settings. Naive Bayes is the basic standard text classifier which is commonly used in practice for text classification [4, 6]. Naïve Bayes proven to works well in particular cases, thus we imply NB in this research to identify the effectiveness with our set of data. Meanwhile, Random Forest (RF) classifiers have the ability of dealing with the high dimensional noisy data in text classification [5] and comprises a set of decision trees each of which is trained using random subsets of feature provides an insight to use this method as a comparison model with SVM and NB.

2 Related Works

This section focuses on related works involving multi-class classification by using SVM, NB and RF in various domain.

2.1 SVM as Multi-class Classifier

Due to SVM ability in constructing linear or nonlinear classification boundaries and produce a promising solution through its support vectors, SVM efficiently classified both binary and multi-class dataset. Burget al. [13] proposed a generalized multi-class SVM to reduce misclassification errors in the loss function and eventually succeed in speed up training time. Rathgamage and Duleep [14] claimed SVM provides less predicting time, less computational complexity while yielded significant prediction accuracy. Based on [15], SVM act as divide and conquer algorithm where it able to reduce number of binary decisions into a multi-class classification decision. Recently, SVM has been used in text classification [16] and successfully reduced the problem of biased when dealing with imbalanced dataset classes. Current research by [17] hybridized SVM with NB to increase the performance of classifier over the TF-IDF vectorization for text categorization.

2.2 NB as Multi-class Classifier

Given that NB is one of ancient method and still relevance with several modification and enhancement mapped to current problem and domain. With the trend of sentiment analysis rising, NB being used as an approach in feature selection for deep learning analysis [18]. NB works well in eliminating mislabel or contradictory instances from training set within multi-class classification [19]. As for the usage of NB in sentiment analysis, [20] proved that NB yielded remarkable accuracy for multi-class classification in semi-supervised learning. Ramesh et al. [21] successfully modified a NB method with a deep learning auto-encoder and GloVe model for feature extraction to serve as an enhanced classifier for multi-class dataset. NB also being used for text classification with multi-class label as [22] successfully proved the efficiency of NB as classifier by integration with associative rule-based.

2.3 RF as Multi-class Classifier

As Random Forest also an excellent classifier when dealing with binary class, facing a multi-class problem required this approach to be enhanced in many ways. Chaudary

et al. [23] claimed by improving RF attribute evaluator method and variable filter method, the accuracy of multi-class classifier by modified RF have increased. RF also widely used as feature extraction as NB. An enhanced RF with multi block local binary pattern by [24] successfully classify multi-class dataset via improving the feature selection method. In comparative research by [25], RF turns out to produced higher accuracy in multi-class classification as compared to other decision tree methods (J48, REPTree and NB) when dealing with medical dataset with variants variables. RF also shown remarkable findings in text classification area where [26] successfully integrate RF with tr-k method to improve the quality and diversity of text features. Other than that RF also proven to produce better accuracy when dealing with imbalance dataset [27] for multi-class classification.

Thus, based on these compiled advantages and abilities of SVM, NB and RF algorithms as multi-class classifier, we perform analysis with additional metrics to evaluate the influence of metric combination toward imbalance datasets.

3 Methodology

This section focuses on data description and phases of methodology involved in brief.

3.1 Data Description

Three datasets based on movie review from three different source are used for this experiment. Table 1 elaborate the data description.

As shown in Table 1, all three datasets comprises of more than two class label. Thus, multi-class algorithm is needed to classify these text data. These datasets gathered from Twitter text specifically on movie reviews for the user to convey their opinion toward the movie and also to recommend the movie based on other users

Table 1 Movie review datasets

Dataset	Details	Class label
A	Movie review from HBO	−1: negative 0: neutral +1: positive
B	Movie recommendation for Netflix viewer	−1: not recommended 0: neutral +1: recommended
C	Movie recommendation For DisneyPlus viewer	−1: not recommended 0: neutral +1: recommended

rating. Based on these text data, we performed sentiment analysis as to categorizing viewer's opinion for the movie's rating.

3.2 Experimental Setup

Figure 1 shows a framework for experimental setup for this analysis.

Data Filtering

All three datasets undergo preprocessing tasks to normalize and reduce redundancies before proceed to text classification. The data are filtered by removing repeating words, removing out of vocabulary words, and correcting misspell words. Cleaned and processed online datasets will produce better classification results rather than blindly classify raw data directly from the online social networks.

SVM Classification

Categorizing movie review by using SVM are as follows:

Step 1. Input: Load training dataset
Step 2. Split the data into word-attribute and class labels

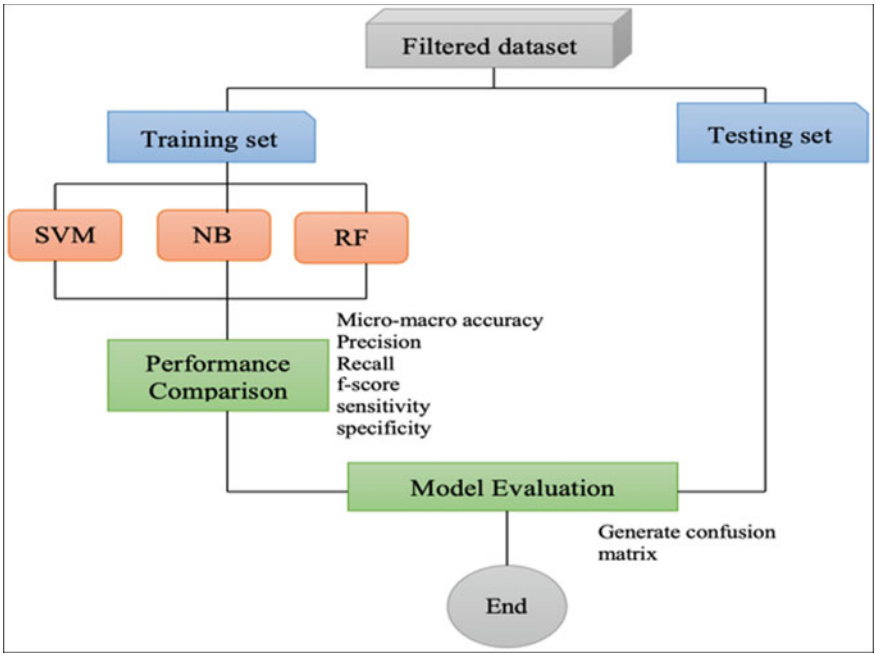


Fig. 1 A framework for experimental setup

Step 3. Construct boundaries to generate classes
Step 4. Construct hyperplane line between classes
Step 5. Calculate similarity distances in words to classify
Step 6. Output: categorized text review.

NB Classification

Categorizing movie review by using NB are as follows:

Step 1. Input: Load training dataset
Step 2. Determine the mean and standard deviation for all variables
Step 3. Perform Gauss density equation to produce probability of all predictor variables in classes
Step 4. Repeat Step 3 until probability of all predicted variables has been calculated
Step 5. Compute the similarity value for each class
Step 6. Output: the greatest similarity value of categorized text review.

RF Classification

Categorizing movie review by using RF are as follows:

Step 1. Input: Load training dataset
Step 2. Predict and store the label of each randomly created decision tree
Step 3. Compute the total votes for individual class (using TDF-IF)
Step 4. Declare majority class as the final categorized class
Step 5. Output: Predicted class of categorized text review.

3.3 Performance Measure

The difference between micro-average and macro-average is the metric for macro-average calculated independently for each class before consider the average of all classes equally, whereas a micro-average compute the average metric by aggregate the contributions of all classes. We use five metrics with regards of macro and micro-accuracy which are precision, recall, f-score, sensitivity and specificity.

$$\text{Precision}_{\text{macro}} = \frac{1}{\text{Class}_{\text{num}}} \sum_{\text{class}} \frac{\text{True Positive}_{\text{class}}}{\text{True Positive}_{\text{class}} + \text{False Positive}_{\text{class}}} \tag{1}$$

$$\text{Precision}_{\text{micro}} = \frac{\Sigma_{\text{class}} \text{True Positive}_{\text{class}}}{\Sigma_{\text{class}} \text{True Positive}_{\text{class}} + \Sigma_{\text{class}} \text{False Positive}_{\text{class}}} \tag{2}$$

$$\text{Recall}_{\text{macro}} = \frac{1}{\text{Class}_{\text{num}}} \sum_{\text{class}} \frac{\text{True Positive}_{\text{class}}}{\text{True Positive}_{\text{class}} + \text{False Negative}_{\text{class}}} \tag{3}$$

$$\text{Recall}_{\text{micro}} = \frac{\Sigma_{\text{class}} \text{True Positive}_{\text{class}}}{\Sigma_{\text{class}} \text{True Positive}_{\text{class}} + \Sigma_{\text{class}} \text{False Negative}_{\text{class}}} \tag{4}$$

$$\text{f-score}_{\text{macro}} = \frac{1}{\frac{1}{\text{Recall}_{\text{macro}}} + \frac{1}{\text{Precision}_{\text{macro}}}} \tag{5}$$

$$\text{f-score}_{\text{micro}} = \frac{1}{\frac{1}{\text{Recall}_{\text{micro}}} + \frac{1}{\text{Precision}_{\text{micro}}}} \tag{6}$$

$$\text{Sensitivity}_{\text{macro}} = 1 - \text{False Negative Rate}_{\text{macro}} \tag{7}$$

$$\text{Sensitivity}_{\text{micro}} = 1 - \text{False Negative Rate}_{\text{micro}} \tag{8}$$

$$\text{Specificity}_{\text{macro}} = 1 - \text{False Positive Rate}_{\text{macro}} \tag{9}$$

$$\text{Specificity}_{\text{micro}} = 1 - \text{False Positive Rate}_{\text{micro}} \tag{10}$$

As for this experiment, we have three datasets with three class label each. Hence, $class_{num}$ is equal to three; ($class_{num} = 3$).

4 Result and Analysis

Table 2 shows the result obtained after dataset A being classified by using three different multi-class algorithms. Result for dataset B and C shown in Tables 3 and 4, respectively.

Based on multi-class classification result for dataset A shown in Table 2, highest average accuracy shown by micro-accuracy in RF algorithm with 88% recorded even highest precision achieved by SVM algorithm by 1.00 for micro precision. High micro-accuracy result also supported by the value of highest f-score (0.88) for

Table 2 Classification results for dataset A

Algorithm	Average Accuracy	Recall	Precision	f-score	Sensitivity	Specificity
SVM	Macro-accuracy	63.50	0.53	0.72	0.61	0.53	0.68
	Micro-Accuracy	83.50	0.67	**1.00**	0.81	0.67	0.85
NB	Macro-accuracy	57.00	0.49	0.63	0.55	0.49	0.61
	Micro-Accuracy	85.50	**0.86**	0.80	0.83	**0.86**	**0.93**
RF	Macro-accuracy	54.75	0.56	0.50	0.53	0.56	0.60
	Micro-Accuracy	**88.00**	0.85	0.91	**0.88**	0.85	0.88

Table 3 Classification results for dataset B

Algorithm	Average Accuracy		Recall	Precision	f-score	Sensitivity	Specificity
SVM	Macro-accuracy	74.50	0.72	0.70	0.71	0.72	0.84
	Micro-Accuracy	**94.50**	**0.93**	**0.98**	**0.95**	**0.93**	**0.95**
NB	Macro-accuracy	66.50	0.76	0.57	0.65	0.76	0.68
	Micro-Accuracy	86.50	0.90	0.95	0.92	0.90	0.90
RF	Macro-accuracy	72.00	0.56	0.90	0.69	0.56	0.71
	Micro-Accuracy	76.50	0.63	0.91	0.74	0.63	0.77

Table 4 Classification results for dataset C

Algorithm	Average Accuracy		Recall	Precision	f-score	Sensitivity	Specificity
SVM	Macro-accuracy	72.50	0.69	**0.76**	0.72	0.69	0.73
	Micro-Accuracy	72.50	0.69	**0.76**	0.72	0.69	0.73
NB	Macro-accuracy	62.25	0.63	0.61	0.61	0.63	0.65
	Micro-Accuracy	62.25	0.63	0.61	0.61	0.63	00.65
RF	Macro-accuracy	**79.50**	**0.85**	0.74	**0.79**	**0.85**	**0.80**
	Micro-Accuracy	**79.50**	**0.85**	0.74	**0.79**	**0.85**	**0.80**

RF algorithm. While NB shown better value of recall, sensitivity and specificity by 0.86, 0.86 and 0.93, respectively.

As for dataset B classification results shown in Table 3, Micro-accuracy for SVM classifier clearly yielded a significant result with highest value for all five metrics of recall (0.93), precision (0.98), f-score (0.95), sensitivity (0.93) and specificity (0.95). These metrics combines and produces higher average micro-accuracy with value of 94.5%.

However for dataset C, RF turns out to produce better accuracy as four metrics combines yielded higher accuracy as compared to others. Four metric which are recall (0.85), f-score (0.79), sensitivity (0.85) and specificity (0.80) contributes to higher average micro-accuracy (79.5%) for dataset C (Fig. 2).

The difference between macro and micro averaging is cleared with macro consider the value of each class equally while micro takes the value of each sample equally. As for dataset C, we have an equal number of samples for each class, then macro and micro resulted in the same accuracy rate. As shown in the classification results for dataset A and B where the imbalance datasets occurred, the value of micro-accuracy

Fig. 2 Comparison graph for macro-accuracy versus micro-accuracy

in average outperform the value of macro-accuracy. We take into consideration the amount of other metrics which contributes the most to the multi-class accuracy result as supported value since a single metric may results in biased and drifting goals. Thus, size of samples in a dataset does influence the accuracy measure for multi-class classification.

5 Conclusion

We proposed multiple metrics to measure the performance of multi-class classifier and successfully analyzed the influence of imbalance datasets toward various classification algorithms. Unlike binary classification, multi-class classification considered as sensitive datasets where the accuracy may easily tampered by the sample size. High precision result may not solely reliable to determine better accuracy when dealing with imbalance dataset. Macro-accuracy majorly used for binary classification as micro-accuracy is usually compute to provide alternatives for determining the accuracy by considering each classes individually. However, not every multi-class dataset can be fixed by using micro-accuracy alone for its metrics. Thus, for further experiment in this analysis can be done by balancing and optimizing the classes before classification takes place.

Acknowledgements We are very grateful to Universiti Teknologi MARA for supporting this research and sincere thanks to our research lab, Data Mining and Optimization (DMO) of Fakulti Teknologi dan Sains Maklumat (FTSM), Universiti Kebangsaan Malaysia, for the expert's knowledge sharing. Massive appreciation to Universiti Kebangsaan Malaysia for providing grant code GGP-2020-032 for this research funding.

References

1. Chaitra PC, Kumar RS (2018) Review of multi-class classification algorithms. Int J Pure Appl Math 118(14):17–26
2. Matthew B, Sohini R (2015) A generalized flow for multi-class and binary classification tasks: an azure ML approach. In: IEEE international conference on big data (Big Data), pp 1728–1737
3. Ghareb AS, Bakar AA, Hamdan AR (2015) Hybrid feature selection based on enhanced genetic algorithm for text. Elsevier J Exp Syst Appl 18:21–44
4. Al-Aidaroos KM, Bakar AA, Othman Z (2010) Naïve Bayes variants in classification learning. In: International conference on information retrieval & knowledge management, IEEE Xplore, pp 276–281
5. Ahmad IS, Bakar AA, Yaakub MR (2020) A survey on machine learning techniques in movie revenue prediction. Springer Nat Comput Sci 1(235)
6. Awwalu J, Bakar AA, Yaakub MR (2019) Hybrid N-gram model using Naïve Bayes for classification of political sentiments on Twitter. Springer Nat Neural Comput Appl 31:9207–9220
7. Yaakub MR, Latiffi MIA, Zaabar LS (2019) A review on sentiment analysis techniques and applications. IOP Conf Ser Mater Sci Eng 55

8. Grandini M, Bagli E, Visani G (2020) Metrics for multi-class classification: an overview. Comput Sci Math. arXiv:221112671
9. Yang Y, Miller C, Jiang P, Moghtaderi A (2020) A case study of multi-class classification with diversified precision recall requirements for query disambiguation. In: Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval, pp 1633–1636
10. Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manage 45(4):427–437
11. Kuo KM, Yalley P, Kao Y, Huang CH (2020) A multi-class classification model for supporting the diagnosis of type II diabetes Melitus. PeerJ 8:e9920
12. Koyejo O, Natarajan N, Ravikumar P, Dhillon IS (2015) Consistent multilabel classification. Comput Sci Math NIPS 7
13. Burg GJJV-D, Groenen PJF (2019) GenSVM: a generalized multiclass support vector machine. J Mach Learn Res 17:1–42
14. Rathgamage D, Duleep PW (2018) Multiclass classification using support vector machines. Electron Theses Dissertations 1845
15. Rathgamage D, Iacob IE (2019) DCSVM: fast multi-class classification using support vector machines. Springer Int J Mach Learn Cybern
16. Saigal P, Khanna V (2020) Multi-category news classification using support vector machine based classifiers. Springer Nat Appl Sci 2(3)
17. Sueno HT, Gerardo BD, Medina RP (2020) Multi-class document classification using support vector machine (SVM) based on improved Naïve Bayes vectorization technique. Int J Adv Trends Comput Sci Eng 9(3):3937–3944
18. Liu Y, Bi JW, Fan ZP (2017) Multi-class sentiment classification: the experimental comparisons of feature selection and machine learning algorithm. J Exp Syst Appl 80:323–339
19. Singh N, Singh P (2019) A novel bagged Naïve Bayes-decision tree approach for multi-class classification problems. J Intell Fuzzy Syst 36:2261–2271
20. Khan AH, Zubair M (2020) Classification of multi-lingual Tweets, into multi-class model using Naïve Bayes and semi-supervised learning. Springer Nat Multimedia Tools Appl 79(3)
21. Ramesh N, Devi GL, Rao KS (2020) A frame work for classification of multi class medical data based on deep learning and Naïve Bayes classification model. Int J Inform Eng Electron Bus 1:37–43
22. Hadi W, Al-Radaideh QA, Alhawari S (2018) Integrating associative rule-based classification with Naïve Bayes for text classification. Appl Soft Comput J 69:344–356
23. Chaudary A, Kolhe S, Kamal R (2016) An improved random forest classifier for multi-class classification. Inform Process Agric 3(4):215–222
24. Kang M-J, Lee J-K, Kang J-W (2017) Combining random forest with multi-block local binary pattern feature selection for multi-class head pose estimation. PLOS One 12(7)
25. Apao NJ, Feliscuzo LS, Romana CLCS, Tagaro JAS (2020) Multiclass classification using random forest algorithm to prognosticate the level of activity of patients with stroke. Int J Sci Technol 9(4):1233–1240
26. Sun Y, Li Y, Zeng Q, Bian Y (2020) Application research of text classification based on random forest algorithm. In: 3rd international conference on advanced electronic materials, computers and software engineering, pp 370–374
27. Tripathi A, Goswami T, Trivedi SK, Sharma RD (2021) A multi class random forest (MRCF) model for classification of small plant peptides. Int J Inform Manage Data Insight 1(2)

A Food Constraint Satisfaction System-Based on Genetic and Random Walk Algorithms

Anilkumar Kothalil Gopalakrishnan

Abstract This paper has presented a novel concept of a food constraint satisfaction system based on the genetic algorithm (GA) and random walk (RW). The system is a part of the food recommendation system under consideration for recommending proper calorie daily food for obese individuals with overweight. The system estimates the values of overweight, daily needed food calorie (DNC), and saturated DNC (DNC_{sat}) of the individuals from their Body Mass Index (BMI) and Basal Metabolic Rate (BMR) values. The daily food set is the food items for breakfast (BF), lunch (LN), and dinner (DN) selected by an individual. Individuals can choose any food item from the predefined food variables for their daily food set. The overall daily food calorie (ODC) is the total food calories estimated from the daily food set. The system aims to assign only the DNC_{sat} matching food items for the daily food set of each individual from the food list of the system. The system has utilized two random-based algorithms such as GA and RW, to find an appropriate daily food set that satisfies the individual's food choices and DNC_{sat} values. The fitness ratio (FR) of GA and RW's node evaluation (ND_{eval}) are estimated based on the individuals' ODC and DNC_{sat} values. The carried-out simulations show that RW outperforms GA in performance and computational complexity. The presented food constraint satisfaction system could effectively satisfy individuals with different food habits and tastes.

Keywords BMI · BMR · Daily needed food calorie · Saturated daily needed food calorie · Food constraint variables · Food dataset · Genetic algorithm · Random walk · Overweight value · Number of days needed

1 Introduction

This research aims to support obese people in solving their food constraint issues by generating healthy and weight-satisfied calorific food items. First, the overweight

A. K. Gopalakrishnan (✉)
Department of Computer Science, Vincent Mary School of Science and Technology, Assumption University of Thailand, Samut Prakan, Thailand
e-mail: anil@scitech.au.edu

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_9

value is estimated from their Body Mass Index (BMI) and Basal Metabolic Rate (BMR). Then, the Daily Needed Calories (DNC) and the saturated DNC (DNC_{sat}) are calculated from the estimated overweight value and the BMR of the person. This paper is just an extension of one of the previously carried out research papers that had described the identification of DNC_{sat} daily food sets without considering any user food habit or choice, especially for vegetarians [1]. As an extension of a previously carried out research, this paper focuses on the satisfactory solutions for the food constraint issues of individuals with a BMI value of 25 or more. From their BMI and DNC_{sat} values, the system would estimate the number of days needed for them to reach an average level of BMI with their constraint satisfied daily food choices, including *breakfast* (BF), *lunch* (LN), and *dinner* (DN). The system allows users to select their favorite BF, LN, and DN food items from a set of predefined food variables. The system defines the food variables based on the two popular food datasets [2, 3].

This paper aims to solve a finite Constraint Satisfaction Problem (CSP) with a limited set of variables and finite domain values. Furthermore, this research has employed the application of two random-based algorithms, genetic algorithm (GA) and random walk (RW), to solve the user food constraints with the most consistent and complete assignments. A Constraint Satisfaction Problem or CSP is solved when each of its variables has satisfied with a deal that meets all the conditions of the variables [4, 5]. In this research, the food constraints of each user are identified and solved by assigning the appropriate food values to their BF, LN, and DN food variables from the food datasets.

2 Review of Related Literature

The Body Mass Index (BMI) is a scientific measure healthy weight status of individuals based on their height and weight values [6]. Large values of this measure indicate various classes of obesity, and individuals with such levels are always at high risk of various body disorders, including diabetes, myeloma, etc. In this paper, the BMI measure acts as a pivot for identifying overweight individuals and estimating their fat values to find the number of days to reach an average BMI value with their constraint satisfied daily food. The food menu selection support system with constraints for a safe dietary life [7] describes how a menu-based system supports individuals with food constraints toward allergic food items. The *k*-NN algorithm was suitable for finding the proper food calories for individuals with higher BMI values [1]. However, it lacks in considering food interests and habits of individuals. A food selection model has solved GA's food supply chain issues [8]. It successfully minimized the total cost under overall quality level logistic constraints, and it has nothing to do with user-level food constraints.

This research is just an extension of a previously carried out research paper [1] that focused only on food recommendations for overweight individuals based on their $DNS_{sat.}$ The concept of user food constraint satisfaction is not a part of that paper. As

the previous research, this study has adopted the ideas of BMI and BMR to estimate the values of overweight and DNC_{sat} of individuals. Although, at present, there are more than 250 equations for BMR are available for investigations [9–12], this paper has utilized one of the most profound BMR equations for the scientific estimation of daily needed body calories (DNC) of individuals with abnormal BMI statuses.

3 The Importance of BMI and BMR in This Research

The BMI value is significant for this research for getting an individual's body weight status and overweight deal. It has nothing to do with any food habit indication of the person [13]. The BMR is the body calorie expenditure indication based on weight, height, age, and gender parameters [13, 14]. Both the values of BMI and BMR are used to estimate the minimum or maximum number of days needed for an abnormal BMI person to gain a "normal" BMI status. The BMI of a person can be estimated as [14]:

$$\text{BMI} = (\text{weight in Kilogram})/(\text{height in meters})^2 \tag{1}$$

Table 1 gives the body weight statuses of individuals as per (1). The overweight value W_{over} of a person is calculated from the height (in meters) and the normal body weight range (boundary) of the BMI (it is from 18.5 to 24.9, see Table 1), which is called BMI_{normal}. The W_{over} (in kilogram) is given as [1]:

$$W_{over} = \text{BMI}_{normal} * \text{height}^2 \tag{2}$$

The upper boundary value of W_{over} is called overweight upper (W_{upper}) and it is calculated from the uppermost value of the BMI_{normal} (24.9). The W_{upper} is given as [1]:

$$W_{upper} = 24.9 * \text{height}^2 \tag{3}$$

Table 1 Body weight status by the BMI

BMI	Body weight status
Less than 15	Very severely underweight
Between 15.0–15.9	Severely underweight
Between 16.0–18.4	Underweight
Between 18.5–24.9	Normal weight
Between 25.0–29.9	Overweight
Between 30.0–34.9	Class I obesity
Between 35.0–39.9	Class II obesity
Above and equal to 40	Class III obesity

From (3), the minimum overweight ($W_{\text{over-min}}$) of a person can be calculated as [1]:

$$W_{\text{over-min}} = (\text{current}_{\text{weight}} - W_{\text{upper}}) \tag{4}$$

Similarly, the maximum overweight ($W_{\text{over-max}}$) of a person is estimated from the lower boundary value of W_{over}, which is called W_{lower}, The value of W_{lower} is calculated from the lowermost boundary value of the $\text{BMI}_{\text{normal}}$ (18.5) and individual's height. The W_{lower} can be calculated as [1]:

$$W_{\text{lower}} = 18.5 * \text{height}^2 \tag{5}$$

From (5), the maximum overweight value ($W_{\text{over-max}}$) of an obese person is calculated from his or her current body weight ($\text{current}_{\text{weight}}$) and W_{lower} [1]:

$$W_{\text{over-max}} = (\text{current}_{\text{weight}} - W_{\text{lower}}) \tag{6}$$

This paper has employed the Harris–Benedict version of BMR calculation for both men and women and is given as [15]:

$$\begin{aligned}\text{Men} =& 88.362 + (13.397 * \text{ weight in kilogram}) \\ &+ (4.799 * \text{ height in centimeter}) - (5.677 * \text{ age in years})\end{aligned} \tag{7}$$

$$\begin{aligned}\text{Women} =& 447.593 + (9.247 * \text{weight in kilogram}) \\ &+ (3.098 * \text{height in centimeter}) - (4.33 * \text{age in years})\end{aligned} \tag{8}$$

Table 2 gives the calculation of the daily needed calorie (DNC) of an individual from daily physical activity and the BMR value [9–12]. For example, consider a person with lightly active daily physical activity has a BMR value of 1924.4 cal. Based on Table 2, the DNC of the person is 2646.05 (1924.4 * 1.375). It is just the daily needed calories to manage their various body functions and life. If the person would like to reduce his weight, then one of the scientific ways is to impose a calorie reduction of 500 cal from the DNC of the person [13, 15]. The person should maintain a total of 2146.05 cal (2646.05 − 500) per day, and it causes a weight loss of 1 pound (or 0.453592 kg) per week, which is equivalent to losing 3500 cal.

Table 2 Estimated values of DNC by BMR

Daily physical activity	DNC
Sedentary or little active	BMR * 1.2
Lightly active	BMR * 1.375
Moderately active	BMR * 1.55
Very active	BMR * 1.725
Extra active	BMR * 1.9

The process of reducing 500 cal from the DNC of a person is called saturation of the DNC (DNC_{sat}), and the process is given as [1]:

$$DNC_{sat} = DNC - 500 \tag{9}$$

From (9), it is possible to conclude that any person who consumes DNC_{sat} level food calories daily will cause a weight reduction of 0.06485 kg per day from their body weight. From (4), the minimum number of days needed (NDN_{min}) for a person to reach the uppermost value of the BMI_{normal} (which is 24.9) is given as [1]:

$$NDN_{min} = W_{over-min}/0.06485 \tag{10}$$

Similarly, from (6), the maximum number of days needed (NDN_{max}) for a person to reach the lowermost boundary value of the BMI_{normal} (which is 18.5) is given as [1]:

$$NDN_{max} = W_{over-max}/0.06485 \tag{11}$$

4 Details of the Food Constraints

This food constraint satisfaction system does not support individuals with a standard or underweight BMI value. This research has classified individuals into two types of food consumers: vegetarian (VG) and non-vegetarian (NV).

There two types of food datasets [2, 3] are used in this research for the creation of the *food list,* and it consists of both the Asian (mainly Indian) and Western food items with their corresponding food variables. At this moment, the size of the *food list* has limited to a set of 500 rows. Each row of the *food list* has filled with a predefined set of food constraint variables and their related food items with their *names*, *sizes,* and *calories***.** Twenty food constraint variables are defined for this research to support individuals with both vegetarian and non-vegetarian food habits. Based on the food variables that a user selects for their BF, LN, and DN, the system can estimate their choices of food constraints. Therefore, it can deliver proper daily food sets for each individual without violating their food constraints. Table 3 shows the details of the twenty food variables defined in the system with their user types and food values.

5 Description of the Constraint Satisfaction Problem

A Constraint Satisfaction Problem (CSP) is formulated with three major components *VER*, *DOM*, and *CS*: (i). *VER* is a set of constraint variables, $\{VER_1, \ldots, VER_n\}$

Table 3 Twenty constraint variables with their food items and user type

User type	Constraint variable	Variable value
VG/NV	M	Milk items such as plain milk, yoghurt, curd, ghee, fruit-milk, smoothie, etc
VG/NV	F	Fruit items such as fresh fruits, fruit juice, fruit salad, etc
VG/NV	O	Onion items such as raw onion/garlic, cooked onion/garlic onion gravy, onion ingredients, onion rings, etc
VG/NV	H	Mushroom items such as fried mushroom, boiled mushrooms, mushroom soup, mushroom noodle, etc
VG/NV	U	Food items with tofu and/or any soy products
VG/NV	T	Black tea, black coffee, milk tea, milk coffee, sugar tea, sugar coffee, etc
VG/NV	A	Any type of alcoholic beverages such as beer, brandy, etc
VG/NV	D	Any type of non-alcoholic beverages such as coke, etc
NV	S	Any type of fish and/or sea food items
NV	G	Egg items such as boiled egg, omelet, or any egg-based food
NV	C	Chicken items such as fried chicken, chicken curry, boiled chicken, chicken rice, etc
NV	**B**	Any type of beef/cow meat food items
NV	P	Any type of pork/ham meat food items
NV	L	Any type of lamb/mutton/goat meat food items
VG/NV	V	Any type of vegetable curry, *veg-biriyani*, vegetable soup, pulav, potato items such as curry, boiled, and fried potatoes
VG/NV	R	Rice items including plain rice, boiled rice, and cereals
VG/NV	Z	Indicates any Indian dishes such as *chapatti*, *puri, plain dosa*, *idli*, *masala dosa*, *uppuma*, etc
VG/NV	I	Indicates cooked vegetable pasta, macaroni, pizza, etc
VG/NV	*W*	Indicates Indian snacks such as *pakora*, *halwa*, *jalebi, kheer*, *rasgula*, *gulabjam*, *samosa*, etc
VG/NV	N	Indicates any nut items such as fried nuts, raw nuts, etc.

(ii). *DOM* is a set of domains, $\{DOM_1, \ldots, DOM_n\}$, for each variable (iii). *CS* is a set of problem-specific constraints defined based on the domain values of the variables [4, 5]. Assume that a vegetarian user (VG) selects the following food constraint variables for their BF: {M, F, V, Z, W, and U}. Based on the food variables, the system identifies the BF food constraints as: "*There is no problem with milk in the BF menu*," "*There is no problem with yogurt or curd in the BF menu*," "*There is no problem with fresh fruits, fruit juice and fruit salad in the BF menu*," "*There is no problem with any vegetable curry or soup in the BF menu*," "*There is no problem with vegetable Biriyani in the BF menu*," "*There is no problem with potato items in the BF menu*," "*There is no problem with chapatti, puri, plain dosa, idli, masala dosa, uppuma in the BF menu*," "*There is no problem with sweet dishes such as halwa, Jalebi, Kheer,*

Rasagula, etc.*, in the BF menu*," *and* "*There is no problem with tofu or soy items in the BF menu.*"

The overall daily food calorie (ODC) is the total calories of the daily food set and is the calories of BF, LN and, DN. The ODC is given as:

$$\mathrm{ODC} = \mathrm{BF}_{\mathrm{calorie}} + \mathrm{LN}_{\mathrm{calorie}} + \mathrm{DN}_{\mathrm{calorie}}, \tag{12}$$

where $\mathrm{BF}_{\mathrm{calorie}}$, $\mathrm{LN}_{\mathrm{calorie}}$, and $\mathrm{DN}_{\mathrm{calorie}}$ are the calories of BF, LN, and DN, respectively. The following sections describe how the GA and RW algorithm support the finding of both the user and $\mathrm{DNC}_{\mathrm{sat}}$ satisfied food set (for BF, LN, and DN) from the *food list*.

6 Application of GA in Solving the User Food Constraints

GA is a kind of evolutionary algorithm based on the famous evolution theory of Charles Darwin. Even though the GAs are computational algorithms and it has nothing to do with natural biological evolution, the algorithm always maintains physical terms such as *chromosome population*, *genes*, *fitness*, *natural selection, crossover*, and *mutation* [4]. Apart from the traditional GA, this research uses a non-binary string of genes for the chromosome population, and each chromosome consists of only three genes (each gene for BF, LN, and DN food items). Similarly, the crossover point of each parent chromosome for crossover is selected based on the index values of the genes of the parent chromosomes.

This research has employed a uniform crossover operation of genes for its offspring chromosome generation. Two similar parent genes from either BF or LN or DN of each chromosome in the population are selected for the crossover. The crossover operation is swapping the sections of the similar genes of each chromosome in the population, and the employed swapping function is flexible. There are around twelve types of gene swapping employed in this research for the crossover operation, and the size of the chromosome population is limited to a value of four. The application of GA in this paper is to excavate the list of both user and DNCsat satisfied daily food items for BF, LN, and DN from the *food list* based on their food constraint variables. Once GA has identified the $\mathrm{DNC}_{\mathrm{sat}}$ and user satisfied food set, the user must follow the "estimated number of days" criterion (10), (11) until to reach a state of healthy BMI status.

7 The Details of Implementation of GA in This Food Constraint Satisfaction Problem

The following steps describe various steps of GA for finding the DNC_{sat} daily food items from the *food list*:

(i) Generate a set of initial chromosome populations randomly based on the food constraint variables for BF, LN, and DN foods. For example, suppose that a VG type user has selected the following food constraint variables for BF: {M, F, U, V, Z, N}, the system labels food variables for the BF gene as: {M_{BF}[index], F_{BF}[index], U_{BF}[index], V_{BF}[index], Z_{BF}[index],N_{BF}[index]}, where 'index' is the row value (from 0 to *n*) of the corresponding food item in the *food list*.

(ii) Next is to generate the three genes for BF, LN, and DN from their corresponding food constraint variables by selecting their index values variables from the *food list*. Thus the genes of BF, LN and DN can be indicated as (M_{BF}[0], F_{BF}[10], U_{BF}[8], V_{BF}[10], Z_{BF}[11], N_{BF}[14]), (M_{LN}[4], H_{LN}[17], D_{LN}[5], V_{LN}[0], R_{LN}[10], I_{LN}[11], W_{LN}[12]), and (F_{DN} [14], H_{DN} [8], V_{DN} [22], Z_{DN} [2], I_{DN} [23], W_{DN} [15], N_{DN} [3]) respectively. This research has tentatively limited the index size of each food constraint variable with a value of twenty five.

(iii) Next is to generate a chromosomes by concatenating the three genes of BF, LN, and DN (see Table 4 for a sample chromosome population).

(iv) Next step is to calculate fitness (FN) of the chromosome population-based on the overall daily calories (ODC) and DNC_{sat} values. The FN of a chromosome is calculated only when its ODC is greater than that of the user's DNC_{sat} value. Otherwise, GA will treat the value of FN as one. The FN is calculated as:

$$FN = 1/(ODC - DNC_{sat}) \tag{13}$$

Table 4 Initial chromosomes with their ODC and FR values

No	Initial chromosome	ODC	FR (%)
1	{(M_{BF}[0], F_{BF}[10], U_{BF}[8], V_{BF}[10], Z_{BF}[11], N_{BF}[14]), (M_{LN}[4], H_{LN}[17], D_{LN}[5], V_{LN}[0], R_{LN}[10], I_{LN}[11], W_{LN}[12]), (F_{DN} [14], H_{DN} [8], V_{DN}[22], Z_{DN}[2], I_{DN}[23], W_{DN} [15], N_{DN} [3])}	2676.33	29.56
2	{(M_{BF}[3], F_{BF}[11], U_{BF}[18], V_{BF}[21], Z_{BF}[1], N_{BF}[4]), (M_{LN}[20], H_{LN}[13], D_{LN}[6], V_{LN}[5], R_{LN}[0], I_{LN}[22], W_{LN}[2]), (F_{DN}[19], H_{DN}[24], V_{DN}[2], Z_{DN} [12], I_{DN} [3], W_{DN} [5], N_{DN} [13])}	2581.47	37.11
3	{(M_{BF}[21], F_{BF}[19], U_{BF}[23], V_{BF}[11], Z_{BF}[17], N_{BF}[0]), (M_{LN}[14], H_{LN}[11], D_{LN}[16], V_{LN}[7], R_{LN}[24], I_{LN}[12], W_{LN}[23]), (F_{DN}[4], H_{DN}[13], V_{DN} [4], Z_{DN} [23], I_{DN} [0], W_{DN} [5], N_{DN} [3])}	3250.85	13.23
4	{(M_{BF}[1], F_{BF}[17], U_{BF}[21], V_{BF}[16], Z_{BF}[7], N_{BF}[2]), (M_{LN}[13], H_{LN}[1], D_{LN}[15], V_{LN}[8], R_{LN}[2], I_{LN}[10], W_{LN}[14]), (F_{DN}[4], H_{DN}[3], V_{DN} [6], Z_{DN} [0], I_{DN} [6], W_{DN} [24], N_{DN} [19])}	2896.34	20.10

Similarly, the fitness ratio (FR) of a chromosome is calculated from the total FN value of the whole population and its FN value. The FR calculation is given as:

$$\mathrm{FR} = (\mathrm{FN} \times 100) / \left(\sum_{i=1}^{n} \mathrm{FN}_i \right), \tag{14}$$

where i is the chromosome number, and the size of chromosome population is four ($n = 4$). A chromosome with an FN value of one is the most fitted chromosome in the population, and its FR is always 100%. Therefore, the termination criterion of GA depends on the FN value of the chromosomes in the population. Table 4 gives the four initial chromosomes with their ODC and FR values.

(v) If none of the FR values of the chromosomes in a population is satisfied, then GA will perform a crossover operation on the genes of the chromosomes to generate offspring chromosomes. The size of the offspring population is always the same as that of the parent population. The crossover is made by exchanging the index values of genes of the corresponding food section (BF, LN, and DN).

(vi) Next is the mutation step of GA, and it can solve the issue of the non-progressive state of the chromosome population of GA to their FR values. The mutation operation randomly changes the index values of the genes of the chromosomes in a population before a crossover.

Table 4 gives the details of an initial population of four chromosomes for a VG user with their ODC (12) and FR (14) values. The $\mathrm{DNC}_{\mathrm{sat}}$ of the user is 2210.43 cal (9). Of the four chromosomes, chromosome no. 2 has shown the top FR value (37.11%), and chromosome no. 3 has the most negligible FR value (13.23%). The FR values of chromosomes are not strong enough to terminate the GA. Therefore, GA has employed its crossover operation for the next iteration.

Table 5 gives the results of ODC and FR of the offspring chromosomes after the first crossover operation. The crossover probability equals 0.75 (the size of chromosome population is four, and each chromosome has three genes). Chromosome no. 2 has shown the top FR value (55.74%), and chromosome no. 4 has shown the lowest FR value (8.65%). Hence, the GA will continue its crossover operation until satisfying the FR value.

Figure 1 shows only the chromosomes with their top FR values until GA has terminated at its nineteenth population (its FR is 95.78% and FN is 1). For this experiment, the mutation probability of chromosomes is tentatively set as 0.1 (one mutation in every ten populations), and in the future, it will be tested with various mutation ratios.

One of the main problems noticed from GA is that its crossover operation is limited to merely the index values of the food variables of each gene of the chromosome population. Furthermore, the computational complexity of GA is increased with the population size, and the size of each gene is proportional to the number of food

Table 5 Crossover results of children chromosomes with their ODC and FR values

No	Children chromosome	ODC	FR (%)
1	{(M_{BF}[1], F_{BF}[19], U_{BF}[18], V_{BF}[11], Z_{BF}[17], N_{BF}[0]), (M_{LN}[20], H_{LN}[13], D_{LN}[16], V_{LN}[7], R_{LN}[0], I_{LN}[10], W_{LN}[23]), (F_{DN}[4], H_{DN}[24], V_{DN} [6], Z_{DN}[12], I_{DN}[3], W_{DN}[24], N_{DN}[13])}	2894.44	11.95
2	{(M_{BF}[21], F_{BF}[17], U_{BF}[8], V_{BF}[10], Z_{BF}[1], N_{BF}[0]), (M_{LN}[4], H_{LN}[1], D_{LN}[5], V_{LN}[7], R_{LN}[0], I_{LN}[11], W_{LN}[14]), (F_{DN}[4], H_{DN}[8], V_{DN}[4], Z_{DN}[2], I_{DN}[23], W_{DN}[5], N_{DN}[13])}	2357.13	55.74
3	{(M_{BF}[3], F_{BF}[10], U_{BF}[21], V_{BF}[10], Z_{BF}[7], N_{BF}[4]), (M_{LN}[4], H_{LN}[1], D_{LN}[6], V_{LN}[5], R_{LN}[10], I_{LN}[11], W_{LN}[2]), (F_{DN}[19], H_{DN} [3], V_{DN}[6], Z_{DN}[12], I_{DN}[0], W_{DN}[15], N_{DN}[19])}	2556.01	23.66
4	{(M_{BF}[0], F_{BF}[11], U_{BF}[18], V_{BF}[21], Z_{BF}[11], N_{BF}[14]), (M_{LN}[14], H_{LN}[17], D_{LN}[16], V_{LN}[8], R_{LN}[24], I_{LN}[10], W_{LN}[2]), (F_{DN}[19], H_{DN}[3], V_{DN}[4], Z_{DN}[2], I_{DN}[23], W_{DN}[5], N_{DN}[13])}	3156.01	8.65

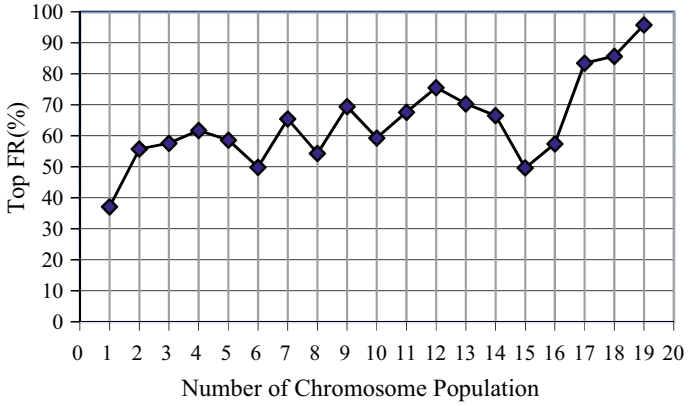


Fig. 1 Top FR values of the populations of GA until its termination

variables from the user. Similarly, the issue of the non-progressive FR states of the chromosome population is unable to solve with the low-level mutation probability of GA, especially in this research.

8 Application of RW in This Food Constraint Problem

In a random walk (RW) algorithm, search happens like a free walk through a set of node points in a problem space until either a node with a satisfying assignment is found [5]. A random walk refers to where one can start from a node that is allocated randomly and is not satisfied with the assignment, then randomly choose a neighbor

node from its node list to navigate a walk until to find a stopping criterion [16–18]. The RW algorithm has been employed in this research with the following steps: (i) Initialize the food constraint variables with their corresponding food details from the *food list*. (ii) Generate a *start node* (called node0) by randomly choosing BF, LN, and DN food constraint indices based on the user's daily food choice (this operation is similar to the generation of genes in GA). Each walking step to the following nodes from the *start node* is assigned an equal probability value. This research has found the ranking of each node is irrelevant. (iii) Instead of ranking of each walked node, this research has computed the evaluation of each walked node (ND_{eval}) and is given as:

$$ND_{eval} = ODC - DNC_{sat} \tag{15}$$

For this paper, the zero or negative value of ND_{eval} is considered the most satisfactory one. (v) If a node with a satisfied ND_{eval}, terminate the walk and display the node details (including the food items, sizes, and calories).

9 Results of RW in the User Food Constraint Problem

Table 6 gives the results of two walk steps by RW algorithm with their node number, ODC value, and ND_{eval} value. The DNC_{sat} of the user is 2430.67 cal. The node0 has a better ND_{eval} (163.16) than the node1 (its ND_{eval} is 219.36).

Tables 7, 8 and 9 give the details of the food items of node0 (in Table 6) including their names, sizes, and calories, respectively. Similarly, the ND_{eval} values of the consecutive nodes generated by RW are shown in Fig. 2. The most promising nodes generated by RW are 7, 12, 13, 16, and 17. At this point, there is no termination criterion for RW is defined.

Table 6 Random walk details from the three nodes (including the start node)

No	Node	ODC	ND_{eval}
0	{(T_{BF}[0], O_{BF}[10], M_{BF}[8], G_{BF}[10], C_{BF}[11], Z_{BF}[14]), (F_{LN}[4], H_{LN}[17], S_{LN}[5], B_{LN}[0], V_{LN}[10], R_{LN}[11]), (A_{DN}[14], O_{DN}[8], S_{DN}[2], B_{DN}[5], V_{DN}[23], Z_{DN}[15], N_{DN} [3])}	2593.83	163.16
1	{(T_{BF}[14], O_{BF}[15], M_{BF}[1], G_{BF}[0], C_{BF}[17], Z_{BF}[8]), (F_{LN}[15], H_{LN}[2], S_{LN}[23], B_{LN}[11], V_{LN}[16], R_{LN}[1]), (A_{DN}[0], O_{DN}[4], S_{DN}[12], B_{DN}[6], V_{DN}[4], Z_{DN}[6], N_{DN} [9])}	2650.03	219.36

Table 7 The food items, size, and calorie of BF constraints T, O, M, G, C, and Z

Constraint variable	Food item	Size	Calorie
$T_{BF}[0]$	Black coffee (no sugar)	1 cup	07.08
$O_{BF}[10]$	Fried onion rings	2 cup	43.92
$M_{BF}[8]$	Milk shake (Sweetened)	10 fl. oz	393.37
$G_{BF}[10]$	Scrambled egg (Oiled)	1 large egg	93.62
$C_{BF}[11]$	Fried chicken with vegetables	1 cup	89.91
$Z_{BF}[14]$	Chapatti	1 Medium	80.00

Table 8 The food items, size, and calorie of LN constraints F, H, S, B, V, and R

Constraint variable	Food item	Size	Calorie
$F_{LN}[4]$	Prune juice (sweetened)	1 cup	181.76
$H_{LN}[6]$	Boiled mushrooms	1 cup	21.84
$S_{LN}[5]$	Fried seafood with vegetables in tomato sauce	1 cup	163.28
$B_{LN}[0]$	Beef noodles in tomato sauce	1 cup	276.39
$V_{LN}[10]$	Boiled mixed vegetables	1 cup	59.15
$R_{LN}[11]$	Fried rice with beans (oiled)	1 cup	195.98

Table 9 The food items, size, and calorie of DN constraints A, O, S, B, V, Z, and N

Constraint variable	Food item	Size	Calorie
$A_{DN}[14]$	Whiskey (scotch)	2 fl. oz	140.0
$O_{DN}[8]$	Garlic bread	2 large slices	135.3
$S_{DN}[2]$	Baked salmon	2_oz	193.914
$B_{DN}[5]$	Beef noodles with vegetables	1 cup	231.57
V_{DN} [23]	Boiled mixed vegetables	1 cup	59.15
$Z_{DN}[15]$	Chapatti	1 Medium	80.0
$N_{DN}[3]$	Fried chickpea	1 cup	147.6

10 Conclusion

The simulations concluded that the employed algorithms are efficient in generating user and DNC_{sat} satisfying daily food set from the *food list*. The presented system has introduced a novel technique for solving food Constraint Satisfaction Problems with a large food dataset. The GA has demonstrated excellent performance in terms of

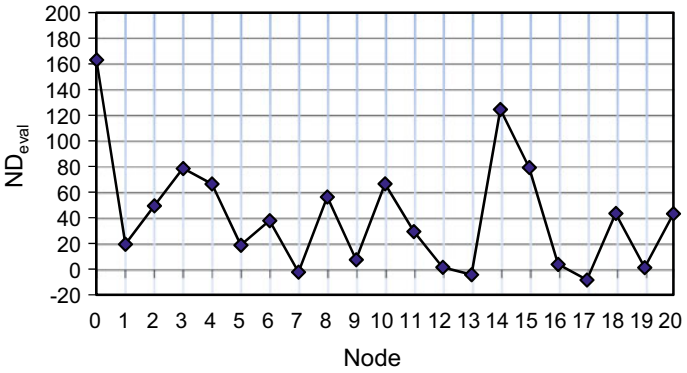


Fig. 2 ND_{eval} values from twenty RW node steps

its searchability despite its computational complexity. Furthermore, the system has strictly instructed its obese users to follow the DNC_{sat} satisfied daily food set until to reach a healthy BMI level. Similarly, the RW approach has shown tremendous performance, and it outperforms the GA in terms of its speed. One of the main contributions of this research is that, it can satisfy the food interests of its users without causing any conflict in their food selection.

References

1. Anilkumar KG (2020) A food recommendation system based on BMI, BMR, *k*-NN algorithm, and a BPNN. In: Proceedings of ICTIS 2020, machine learning for predictive analysis, Springer LNNS book series, vol 141, pp 107–118. https://doi.org/10.1007/978-981-15-7106-0_11(2020)
2. https://catalog.data.gov/dataset/mypyramid-food-raw-data. Last accessed on 22 Nov 2021
3. https://www.nriol.com/health/calorie-chart.asp. Last accessed on 22 Nov 2021
4. Stuartl R, Peter N (2021) Artificial intelligence: a modern approach, 4th edn. Pearson
5. David LP, Mackworth AK (2018) Artificial intelligence foundations of computational agents, 2nd edn. Cambridge University press
6. https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmicalc.htm. Last accessed on 22 Nov 2021
7. Kayo I, Takuya O, Kouki M, Sui K, Yasuki I (2012) Food menu selection support system: considering constraint conditions for safe dietary life. https://doi.org/10.1145/2390776.2390786
8. Douiri L, Abdelouahhab J, Abdellah EB, Moumen AD (2021) Optimization of fresh food distribution route using genetic algorithm with the best selection technique. In: Constraint handling in meta-heuristics and applications. Springer Singapore
9. Sabounchi NS, Rahmandad H, Ammerman A (2013) Best fitting prediction equations for basal metabolic rate: informing obesity interventions in diverse populations. Int J Obes 37(10):1364–1370
10. Abdel-Hamid TK (2002) Modeling the dynamics of human energy regulations and its implications for obesity treatment. Syst Dyn Rev 18(4):431–471
11. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/basal-metabolic-rate. Last accessed on 18 Nov 2021

12. https://www.sciencedirect.com/topics/food-science/food-choice. Last accessed on 20 Nov 2021
13. Rothman KJ (2008) BMI-related errors in the measurement of obesity. Int J Obes 32:56–59
14. https://www.sharecare.com/group/sharecare-fitness. Last accessed on 21 Nov 2021
15. https://www.garnethealth.org/news/basal-metabolic-rate-calculator. Last accessed on 22 Nov 2021
16. https://medium.com/@gbrnc28/random-walk-method-page-rank-algorithm-using-networkx-7da8696ecc38. Last accessed on 5 Nov 2021
17. https://www.britannica.com/science/random-walk. Last accessed on 5 Nov 2021
18. https://machinelearningmastery.com/gentle-introduction-random-walk-times-series-forecasting-python/. Last accessed on 10 Nov 2021

An Ontological Model for Fire Evacuation Route Recommendation in Buildings

Joaquim Neto, A. Jorge Morais, Ramiro Gonçalves, and António Leça Coelho

Abstract The study of the evacuation of buildings in emergency fire situations has deserved the attention of researchers for decades, particularly regarding the real-time guiding of occupants in their way to exit the building. However, finding solutions to guide the occupants evacuating a building requires a thorough knowledge of that domain. Using ontological models to model the knowledge of a domain allows the understanding of that domain to be shared. This paper presents an ontological model that pretends to reinforce and deepen knowledge of the domain under study and help develop solutions and systems capable of guiding the occupants during a building evacuation. The ontology was developed following the *METHONTOLOGY* methodology, and for implementation, the *Protégé* tool was used. The ontological model was successfully submitted to a thorough evaluation process and is publicly available on the Web.

Keywords Ontological model · Knowledge representation · Fire route recommendation · Ontology · Internet of things

J. Neto (✉) · A. Jorge Morais
Universidade Aberta, Lisboa, Portugal
e-mail: jfn@lnec.pt

A. Jorge Morais
e-mail: jorge.Morais@uab.pt

J. Neto · A. L. Coelho
Laboratório Nacional de Engenharia Civil, Lisboa, Portugal
e-mail: alcoelho@lnec.pt

A. Jorge Morais
LIAAD - INESC TEC, Porto, Portugal

R. Gonçalves
UTAD (Universidade de Trás-Os-Montes E Alto Douro), Vila Real, Portugal
e-mail: ramiro@utad.pt

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_10

1 Introduction

The development of solutions capable of guiding, in real-time, the building occupants on their way to exit the building or reach a safe zone in a fire emergency is a problem that deserves the attention of the researchers. The real-time recommendation of evacuation routes is an approach to be considered and studied. However, developing those solutions requires in-depth knowledge of the building evacuation domain, which we can achieve by developing an ontological model. The motivation for this research work falls within the scope of our PhD research work, which aims to propose and evaluate how a multi-agent-based recommender system can guide the occupants in their way to exit a building [1]. The system supports its development on an ontological model that considers the different factors impacting an evacuation route recommendation system, allowing a thorough knowledge of the domain. To build the ontology, we chose to divide our research work into two steps: First, we developed an ontology for the evacuation of buildings under fire emergency [2]. Then, in a second step, and based on the referred ontology, we develop the ontological model presented here using [2] as the basis for the work.

Section 2 presents research works that use ontologies to represent knowledge in the fire emergency research area. Section 3 introduces the concept and how to develop an ontology. Section 4 describes all the steps taken in developing the ontology, and Sect. 5 refers to how the ontology was evaluated. The paper's last section presents the main conclusions, contributions, and plans for future research in this area.

2 Ontologies and Building Emergency: Related Work

Researchers commonly use ontologies for modelling domain knowledge, and the building emergency domain is not an exception. In our literature review, we found many research works and here, we will highlight the more relevant ones. The ontology SEMA4A [3] provides the necessary emergency information to support the interoperability between people and between systems for emergencies, accessibility guidelines, and communication technologies. Using SEMA4A ontology, [4] developed an extended ontology that deals with the notification of evacuees about evacuation routes, exits, and safe places. Liu et al. [5] presents an ontology that considers fire control concepts and their relationships from the community members perspective, defining how they can contribute to controlling fire. Hristoskova et al. [6] presents an ontology focussed on the emergency response that deals with intelligent emergency applications, focussing on interpreting, and filtering relevant information from a large amount of heterogeneous data. The *emergency elements* (EMERGEL) ontology, developed by Casado et al. [7], provides concepts and knowledge related to the emergency area. In their work, [8] presents an ontology aiming to share knowledge in the area of fire emergency response in buildings, helping organisations develop their strategic and tactical plans. Neto et al. [2] presents an ontology that contributes to

a better understanding of the fire building evacuation domain, aiming to help the development of building evacuation systems.

The literature review identifies different ontological approaches to deal with the problem of emergency fire in buildings. Some of the approaches focus on fire control knowledge; others are concerned with representing fire safety knowledge, as is the case with occupants evacuation. However, we did not identify research works focussed on developing ontologies whose focus is the recommendation of evacuation routes in real-time, which is the domain and scope of this research work.

3 Building Ontologies

According to Gruber [9], in the area of information science and computing, the term ontology usually refers to an artefact that models the knowledge of a domain. In the artificial intelligence (AI) literature, [10] refers to the existence of multiple and contradictory definitions for ontology. In the AI context, [11] states that an ontology is an "*explicit specification of a conceptualisation*", where concepts, objects and other entities, and their relationships represent the knowledge of a domain. The ontologies contribute to understanding a domain and help the interoperability between people and heterogeneous systems. An ontology consists of terms representing hierarchically organised concepts and some specifications of their meaning [12].

For ontology development, the most representative and commonly used methodologies [12] are ENTERPRISE [13], TOVE [14], METHONTOLOGY [15], and "Ontology Development 101" [10]. As referred by Fernández-López et al. [15], Pinto and Martins [12], independently of the methodology, the development life cycle typically follows a five stages process. First, the development process starts with the ***specification stage***, where we identify and define the scope and goal of the ontology. Then, a conceptual model is built to describe the ontology according to the specification defined before in the ***conceptualisation stage***. Third, the ***formalisation* stage** is where the conceptual model gives rise to a formal model. Fourth, the ontology is represented through a knowledge representation language at the ***implementation stage***. Finally, the last stage is ***maintenance***, which includes updating and correcting the implemented ontology. During the whole development ontology life cycle, three other activities must be considered [16, 13]: strengthen domain ***knowledge acquisition***, namely supported on the relevant bibliography or interviewing domain specialists; ontology ***evaluation;*** and ontology ***documentation.***

4 Building the Ontology

We considered the METHONTOLOGY methodology for ontology development, following those mentioned above four main steps. In addition, there was a systematic work of study throughout the development process to consolidate knowledge about the domain. Furthermore, the ontology development was also complemented with the documentation and evaluation in the different stages of the process. Therefore, we divide our research work into two phases. First, we start by developing an ontology for the evacuation of buildings under fire emergency [2], and then, we develop the ontological model presented in this paper, reusing that ontology. The ontology presented in Neto et al. [2] aims to develop a knowledge representation about the thematic of buildings evacuation under fire, with the purpose that the ontology may contribute to a better understanding of the domain and help support the development of other ontologies related to the building evacuation domain, such as the one presented here.

4.1 Specification Stage

The ontology's domain, scope, and purpose are defined at the specification stage. According to Neto et al. [2], four main questions need to be answered: ***What is the ontology domain? Why build the ontology? What is the expected use of the ontology? What are the expected answers the ontology should give?***

As in Neto et al. [2], we will use a set of competency questions to help us to define the ontology's specifications and requirements. The competency questions were formulated based on Portuguese regulations and legislation[1] and in a report [16] about the self-protective fire safety measures in buildings and the field experts contribution. Table 1 presents the competency questions and the essential aspects of the answers. Since the present ontology is based on the ontology presented in Neto et al. [2], to a certain extent, it also inherits the competency questions then considered.

From the above, the ontology domain fits the real-time recommendation of evacuation routes to the occupants of a building under fire emergency. As for the second question: Why build the ontology? This research work aims to build a representation model that addresses the real-time recommendation of evacuation routes to occupants in buildings supported on IoT. Concerning the expected use, this ontological model inherits from Neto et al. [2] the ability to strengthen and consolidate knowledge about building evacuation under fire emergencies. Furthermore, it contributes with the knowledge to support the development of evacuation solutions capable of guiding the occupants of a building under fire emergency, such as the one we are developing and studying in the context of our PhD research work [1]. Regarding

[1] Decreto-Lei n.º 220/2008 de 12/11 - Regime Jurídico da Segurança Contra Incêndios em Edifícios (SCIE).

Table 1 Set of competency questions and respective answers

Competency questions	
Q1	*How is the alarm transmitted to the occupants?*
A1	The occupants are notified through visual or audible alarms and messages previously recorded
Q2	*How do the occupants behave in a fire emergency?*
A2	The occupant's behaviour depends on personal characteristics and building knowledge. They also tend to follow the signage
Q3	*How do occupants know the location of building exits and safe zones?*
A3	Emergency signs identify the building's exits and safety zones and their routes
Q4	*How to evacuate the building? How to identify evacuation routes?*
A4	The evacuation process must follow the building's organisation and management. The emergency signs help the occupants exit the building
Q5	*What type of emergency signage exists in the building? How could the emergency signage help occupants throughout the evacuation process?*
A5	Emergency signs are generally static and do not change during the evacuation process. However, dynamic emergency signage updated in real-time provides the occupants with better information about the safest evacuation routes
Q6	*What types of hazards are faced by the building occupants, and how do they influence the building evacuation process?*
A6	The hazards are toxic gases, smoke, or route congestion. They influence the context in which the evacuation takes place, leading to blockage, and congestion of the evacuation routes
Q7	*How can context be captured and used to help occupants?*
A7	An occupant perceives the context that surrounds him. Context is captured by sensors installed in the building and used to help occupants get to a safe location
Q8	*How can context influence the evacuation process?*
A8	The context affects the evacuation routes and impacts how to evacuate the building
Q9	*How can the data produced by the sensors help the occupants in the building evacuation process?*
A9	The sensor's data can be used to notify occupants and support an information system to help guide occupants through the building evacuation process
Q10	*How can the Internet of things (IoT) contribute to creating a solution capable of guiding occupants to a safe location in an emergency fire situation?*
A10	Sensors and digital signage can be IoT devices. The data collected by the IoT input devices (sensors, fire detection systems, smartphones) are sent to a central system for processing. Finally, the processed data are presented (recommended) to the occupants through IoT output devices (digital signage or smartphones)

the fourth question, about the ontology's answers, we must consider the competency questions in Table 1. Besides helping to define the questions that the ontology must be able to answer, the set of competency questions will also be used to test the ontology at the evaluation stage. In Table 2, we summarise the specifications and requirements of the ontology.

Table 2 Ontology specification and requirements

Ontology specifications and requirements	
Domain:	The domain is the real-time recommendation of safe evacuation routes to the occupants of a building under a fire emergency
Goal:	To develop a representation model that addresses the real-time recommendation of evacuation routes to occupants in buildings supported on IoT
Contributions:	Provides the ability to strengthen and consolidate knowledge about building evacuation under fire emergencies Provides a knowledge model to support the development of evacuation solutions to real-time guiding the occupants of a building in a fire emergency
Expected answers:	• How does context influence the occupant behaviour? • How can contextual information be used to guide occupants? • How does the type of emergency signage influence the evacuation process? • How is it possible to provide real-time information to occupants? • How to guide the occupants to exit the building? • How is the building represented on the recommender system? • How are evacuation route recommendations generated?

4.2 Conceptualisation Stage

The knowledge acquired in the specification stage is used to describe the ontological model through a conceptual model constituted by a set of concepts and their relationships. The herein presented conceptual map was built with the *mindomo*[2] framework, on top of the "*conceptual map of the fire building evacuation ontology* "[17], which was the basis of the ontology presented on [2]. Furthermore, based on the ontology's specification and requirements (Table 2) and the answers to the competency questions summarised in Table 1, we identify the following core terms: *route recommender system, Internet of things* (*IoT*)*, IoT device*, and we build the conceptual map [18] presented on Fig. 1. The conceptual map shows the use of the conceptual map presented on [17], highlighting the terms to which the new conceptual map relates.

4.3 Formalisation and Implementation Stages

At this phase of the development process, the conceptual model is transformed into a formal model implemented in a knowledge representation language, as is the case of *Protégé* [19], an open-source platform to create ontologies. With *Protégé,* we create and describe the14 new classes, highlighted in Fig. 2, and 20 object properties

[2] www.mindomo.com

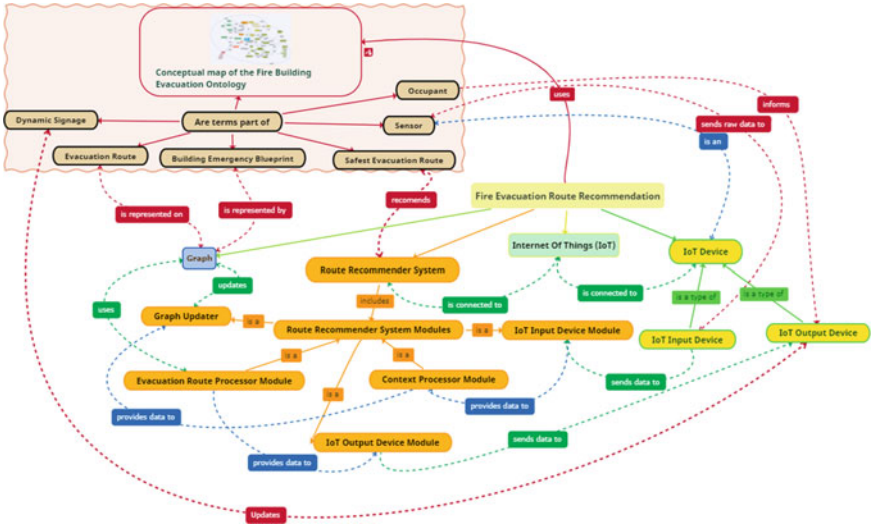


Fig. 1 Conceptual map

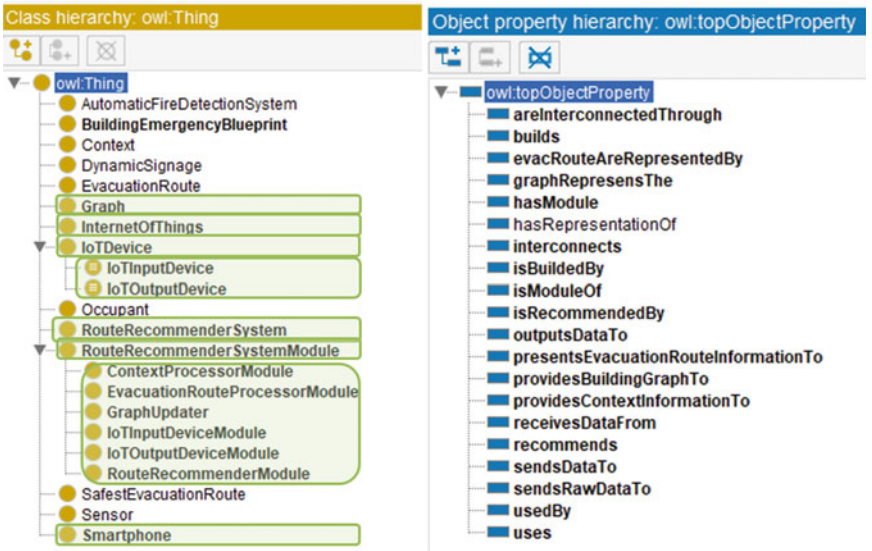


Fig. 2 List of classes and sub-classes

to characterise the relationships of the concepts identified in the conceptualisation stage.

In Fig. 3, we present a *Protégé* screen capture exemplifying the class characterisation with the *RouteRecommenderSystem class.*

Fig. 3 Class representation on *Protégé*

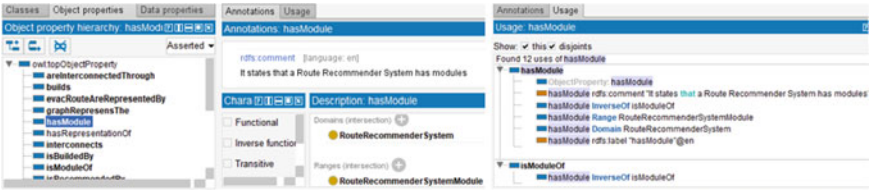


Fig. 4 Object property *hasModule*

Object properties set the relationship between instances of two classes. In Fig. 4, the *hasModule* object property is presented. It shows the relationship between individuals of the classes *RouteRecomenderSystem* (domain) and *RouteRecommenderSystemModule* (range), establishing that a route recommender system consists of modules represented by instances of *RouteRecommederSystemModule* sub-classes.

Datatype properties refer to a relationship of an individual of a class with a primitive value. For example, Fig. 5 shows the property *isAnIoTDevice,* relating an individual of the inferred class (*DynamicSignage* ***or*** *Smartphone* ***or*** *AutomaticFireDetectionSystem* ***or*** *Sensor*) to a primitive of type *Boolean*. Furthermore, the property states that an *IoT* device is any individual of that class if the property *isAnIoTDevice* equals *true.*

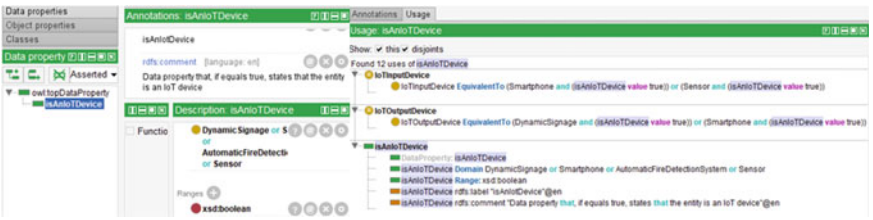


Fig. 5 The *isAnIoTDevice* data property

5 Evaluation

Evaluating an ontology is an iterative process during all the ontology's development life cycle. It is a judgement of the ontology content against a referential [20] that, in this case, is embodied by the ontology requirements and specifications (Table 2) and by the competency questions (Table 1). The technical evaluation considers the aspects of ontology structure and architecture, namely syntactic class and properties validation [2]. In addition, it also deals with ontology documentation to provide an adequate understanding of ontology. Our evaluation process used the tools and services provided by the *Ontology Improvement Tool* (*V2*)[3] Web tool, for ontology consistency, with the *OOps!* (*OntOlogy Pitfall Scanner*)[4] [21], ontology improvement and validation, such as syntactic validation checking, with the RDF *Triple-Checker*,[5] or to verify if the created semantic Web follows best practices,[6,7,8] a service provided by *Vapour*.[9] The tests carried were completed and are available on the ontology URL.[10] To document the ontology, we used *WIDOCO* [22]. This wizard identifies missing metadata and integrates other tools, like *OOPS*! for ontology validation, *LODE*[11] for ontology documentation and *WebVowl*[12] for ontology visualisation. We also evaluate the ontology against the competency questions and the ontology's specifications and requirements to assess whether the ontology classes, properties, and axioms fulfil the ontology requirements. That can be done by querying the ontology using the *SPARQL*. Table 3 shows an example, together with a graphical visualisation provided by *Protégé*.

6 Conclusion and Future Work

The ontological model proposed in this paper aims to make contributions to understanding the domain of real-time recommendation of safe evacuation routes to the occupants of a building under a fire emergency. The main purpose of this ontology is to develop a knowledge model capable of supporting the future development of more suitable building evacuation solutions and systems. The ontology was implemented with the Protégé tool and was successfully evaluated using a set of Web tools and

[3] http://perfectsemanticweb.appspot.com/?p=ontologyValidation.

[4] OOPS!—OntOlogy Pitfall Scanner! (linkeddata.es).

[5] http://graphite.ecs.soton.ac.uk/checker/

[6] Linked Data—Design Issues (w3.org).

[7] Best Practice Recipes for Publishing RDF Vocabularies (w3.org).

[8] Cool URIs for the Semantic Web (w3.org).

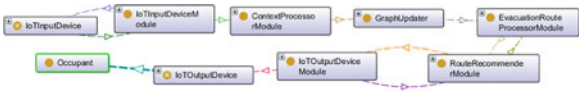
[9] http://linkeddata.uriburner.com:8000/vapour.

[10] https://www.1000palavras.pt/ontology/fEvacRouteRecomm/FireEvacRouteRecommBuilding-en.html.

[11] https://essepuntato.it/lode/

[12] http://vowl.visualdataweb.org/webvowl.html.

Table 3 Example of an answer to a question, SPARQL axioms, and graphical representation

Question	How to guide the occupants to exit the building?
Axioms	*IoTInputDevice* => *sendsRawDataTo* => *IoTInputDeviceModule* => *sendsDataTo* => *ContextProcessorModule* => *providesContextInformationTo* => *GraphUpdater* => *providesBuildingGraphTo* => *EvacuationRouteProcessorModule* => *sendsSafestRoutesTo* => *RouteRecommenderModule* => *providesRecommendationsTo* => *IoTOutputDeviceModule* => *outputsDataTo* => *IoTOutputDevice* => *presentsEvacuationRouteInformationTo* => *Occupant*
Graphical representation	

by querying the ontology with the SPARQL language. The ontology is available at the URI: https://www.1000palavras.pt/ontology/fEvacRouteRecomm/FireEvacRouteRecommBuilding-en. As future work, we intend to deepen the ontology by creating use cases to consolidate the knowledge model and support the development of the multi-agent recommender system presented by Neto et al. [1].

References

1. Neto J, Morais AJ, Gonçalves R, Leça Coelho A (2019) A multi-agent system for recommending fire evacuation routes in buildings, based on context and IoT. In: Communications in computer and information science, pp 343–347
2. Neto J, Jorge Morais A, Gonçalves R, Coelho AL (2022) An ontology for fire building evacuation. Springer, Singapore, pp 975–985
3. Malizia A, Onorati T, Diaz P, Aedo I, Astorga-Paliza F (2010) SEMA4A: An ontology for emergency notification systems accessibility. Expert Syst Appl 37(4):3380–3391
4. Onorati T, Malizia A, Diaz P, Aedo I (2014) Modeling an ontology on accessible evacuation routes for emergencies. Expert Syst Appl 41(16):7124–7134
5. Liu G, Xu B, Tu Q, Sha Y, Xu Z (2011) Towards building ontology for community-based fire management. In: Proceedings 2011 international conference on transportation, mechanical, and electrical engineering, TMEE 2011, pp 1366–1369
6. Hristoskova A, Ongenae F, De Turck F (2013) Semantic reasoning for intelligent emergency response applications. In: IEEE International conference on industrial informatics (INDIN), pp 547–554
7. Casado R, Rubiera E, Sacristan M, Schütte F, Peters R (2015) Data interoperability software solution for emergency reaction in the Europe Union. Nat Hazards Earth Syst Sci 15(7):1563–1576
8. Bitencourt K, Durão F, Mendonça M (2015) EmergencyFire: an ontology for fire emergency situations. In: Proceedings of the 21st Brazilian symposium on multimedia and the web—WebMedia, vol 15, pp 73–76
9. Gruber T (2009) Ontology. In: LIU L, ÖZSU MT (eds) Encyclopedia of database systems. Springer US, Boston, MA, pp 1963–1965
10. Noy N, McGuinness D (2001) Ontology development 101: a guide to creating your first ontology. Protege 32:1–25
11. Gruber T (1995) Toward principles for the design of ontologies used for knowledge sharing. Int J Hum—Comput Stud 43(5–6):907–928

12. Pinto HS, Martins JP (2004) Ontologies: how can they be built? Knowl Inf Syst 6(4):441–464
13. Uschold M, King M (1995) Towards a methodology for building ontologies
14. Grüninger M, Fox MS (1995) Methodology for the design and evaluation of ontologies
15. Fernández-López M, Gomez-Perez A, Juristo N (1997) Methontology: from ontological art towards ontological engineering. Eng Work Ontol Eng
16. Pinheiro J (2012) MEDIDAS DE AUTOPROTEÇÃO DE SEGURANÇA CONTRA INCÊNDIO EM EDIFÍCIOS. Autoridade Nacional de Protecção Civil
17. Neto J (2021) Conceptual map of the fire building evacuation ontology, 2020. [Online]. Available: https://www.mindomo.com/pt/mindmap/38f6214ddb874641b486f3fdc2a06f4a. Accessed 18 Jan 2021
18. Neto J (2021) Fire evacuation route recommeendation 2021. [Online]. Available: https://www.mindomo.com/pt/mindmap/fire-evacuation-route-recommendation-2198e2da36074dd2939793a4a35f7cb4. Accessed 22 Nov 2021
19. Musen MA (2015) The Protégé project: a look back and a look forward. AI Matters 1(4):4–12
20. Gómez-Pérez A, Fernández-López M, Corcho O (2004) Ontological engineering : with examples from the areas of knowledge management, e-commerce and the Semantic Web. Springer
21. Poveda-Villalón M, Gómez-Pérez A, Suárez-Figueroa MC (2014) OOPS! (OntOlogy Pitfall Scanner!): an on-line tool for ontology evaluation. Int J Semant Web Inf Syst 10(2):7–34
22. Garijo D (2017) WIDOCO: a wizard for documenting ontologies. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol 10588. LNCS, pp 94–102

Securing International Space Station Against Recent Cyber Threats

Samaneh Pazouki and Abdullah Aydeger

Abstract Artificial satellites, as a vital part of our infrastructures, are launching to space for different purposes such as the Internet, forecasting weather/disaster, and space exploration. International Space Station (ISS), as a space exploration research center in the Earth orbit, is the largest satellite in space. Astronauts in ISS do research on astrobiology, astronomy, and testing spacecraft equipment for a long-term mission to Moon and Mars. ISS is considered a unique satellite to provide some valuable information about space exploration; however, according to research, the ISS is vulnerable to different cyberattacks. In other words, cybersecurity challenges may compromise the ISS performance, which is controlled via ground/space stations and communication between them. Thus, this paper considers cyberattack threats compromising the ISS and mathematically models false data injection (FDI) attack threatening ISS power system. Two types of FDI are applied to the ISS power system, which results in (1) battery depletion and (2) load shedding. Then, security defense mechanisms are recommended to protect the ISS power system from the cyber threats.

Keywords Cyberattacks · Cybersecurity · Cyber-physical space systems · Defense mechanisms · False data injection cyberattack · International Space Station

1 Introduction

As a vital part of our infrastructures, artificial satellites are exponentially launched to space with different purposes/applications. Players such as SpaceX [1], Blue Origin [2], and Virgin Galactic [3] are investing to X*100 billion $ projects to Mars and Moon. International Space Station (ISS) [4] as a research laboratory in space is

S. Pazouki (✉)
School of Electrical, Computer, and Biomedical Engineering, Southern Illinois University, Carbondale, IL, USA
e-mail: samaneh.pazouki@siu.edu

A. Aydeger
School of Computing, Southern Illinois University, Carbondale, IL, USA
e-mail: aydeger@cs.siu.edu

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_11

orbiting the Earth for different purposes such as GPS, Internet, weather/disaster forecasting, cyber-physical power system, and space exploration [5]. There are cyberattacks that threaten the satellites [6–8]. A secure architecture for satellites is presented in [9]. Some security solutions for NASA center are stated in [10]. Some security solutions for space communication are elaborated in [11]. Some security requirements for autonomous space systems are presented in [12]. The authors in [13] consider space exploration mission and present some relevant cyber-physical challenges. Reference [14] designs a model for the autonomous space vehicle resupplying the ISS.

Scrutinizing previous researches unfold the gap in the existing literature considering cybersecurity challenges/solutions of satellites. *A few* security challenges and defense mechanisms for satellites have been elaborated in the available research; however, *more* research is required to consider cybersecurity challenges and cyber defense mechanisms against the cyber-physical space systems threats. Considering cybersecurity challenges against the ISS, this paper proposes a mathematical model for a FDI cyberattack threat against significant components of each satellite, especially ISS power system, which has not been modeled/considered yet. Two types of FDI cyberattack are presented in this research: (1) FDI resulting in battery depletion of ISS power system and (2) FDI resulting in load shedding of the power system. Considering the proposed FDI cyberattacks impacts, some security defense mechanisms/solutions are recommended in this paper to protect the ISS power system from the threats.

2 Framework and Possible Cyberattacks

2.1 ISS Framework

According to Fig. 1 [15], the ISS has a set of solar panels [16] to power its electrical devices using DC power. The power from solar panels can feed devices in main habitable laboratories such as laptops directly or can be stored in the batteries. A battery charge/discharge unit before the battery [17], based on the commands it receives from the management unit or Earth station, determines the battery must be charged or discharged in each hour. DC switching unit based on a predefined algorithm defines the loads that must be fed by solar panels or batteries at each hour. The switching unit determines the amount of power distributed between three habitable modules. The ISS is also managed by commands received from the Earth station. ISS power system is very important because it is the only source of energy to supply the ISS demands.

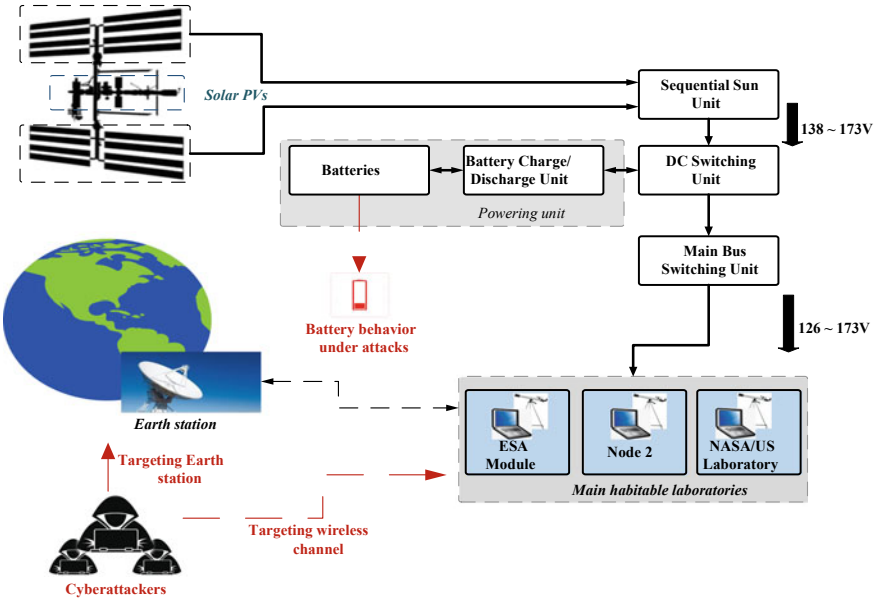


Fig. 1 Framework of ISS power system under cyberattacks

2.2 *Potential Cyberattacks Against the ISS*

The ISS structure has some vulnerabilities [7], which are discussed as follows:

1. The Earth station can be hacked by attackers. The attackers may gain access to sensitive information in the Earth station, send malicious stimulation to ISS via impersonation, and apply Flooding attack to the ISS.
2. Wireless channel between Earth station and ISS might be hacked. Attackers can use this channel to enter the ISS internal system, crash the equipment, stop communicating with the Earth station, and apply jamming attack for channel congestion to affect availability of the ISS.
3. Space tourism may also carry digital devices for entrancing the ISS system and hack the ISS by insider attacks. They can pass the authentication inside the ISS and inject the malicious commands.
4. In addition to NASA, other manufacturers are investing in space projects. Hence, it increases the ways that malicious attackers can unlock the authentication and access the authorization to penetrate the ISS.

3 Mathematical Formulation

3.1 Objective Function

From the attacker viewpoint, the FDI cyberattack should be maximized to significantly affect the ISS power system in (1) depleting the battery in Eq. (1) and (2) load shedding in Eq. (2).

$$\mathrm{Max} I^{\mathrm{FDI}} = \sum_{t=1}^{24} P_t^{\mathrm{PD}} \tag{1}$$

$$\mathrm{Max} I^{\mathrm{FDI}} = \sum_{t=1}^{24} P_t^{\mathrm{LS}} \tag{2}$$

3.2 ISS Power System

The internal electrical system of ISS has three main parts which are described:

3.2.1 Power Balance

Similar to any electrical network, the generated and consumed power in ISS should be equal at each hour in Eqs. (3) and (4). There are some loads which are totally stated P_t^L which is consumed in main habitable module (Fig. 1), while the amount of solar panels power is P_t^{Solar}. The charge/discharge power of battery is stated with P_t^{DisCh} and P_t^{Ch}, respectively. In Eq. (3), P_t^{PD} increases the amount of load from the standpoint of power source to model the power depletion. But in Eq. (4), P_t^{LS} is decreasing the amount of load to model load shedding.

$$P_t^L + P_t^{\mathrm{PD}} = P_t^{Solar} + P_t^{DisCh} - P_t^{Ch} \tag{3}$$

$$P_t^L - P_t^{\mathrm{LS}} = P_t^{Solar} + P_t^{DisCh} - P_t^{Ch} \tag{4}$$

3.2.2 Battery

ISS is supplied through power of solar panels or batteries. Batteries can be charged or discharged, which are switched by DC switching unit (shown in Fig. 1) based

on commands received from internal management unit of ISS or Earth station. According to Eq. (5), the batteries used in ISS must contain a suitable level of energy at first hour of each day ($B_{t=0}^{Batt}$) to afford the load supplement. In addition to existing lower ($\underline{B}^{Batt}$) and upper ($\overline{B}^{Batt}$) limits of energy capacity defined by Eq. (6), the power through discharge and charge process also has an upper limit ($\overline{P}^{Batt}$), which is defined by Eqs. (7) and (8), respectively. Equation (9) determines how battery energy (B_t^{Batt}) is associated with charging and discharging power.

$$B_{t=0}^{Batt} = P_t^L \tag{5}$$

$$\underline{B}^{Batt} \leq B_t^{Batt} \leq \overline{B}^{Batt} \tag{6}$$

$$0 \leq P_t^{DisCh} \leq \overline{P}^{Load} \tag{7}$$

$$0 \leq P_t^{Ch} \leq \overline{P}^{\mathrm{PV}} \tag{8}$$

$$B_t^{Batt} = B_{t-1}^{Batt} - \left(1/\phi^{DisCh}\right) P_t^{DisCh} + \phi^{Ch} P_t^{Ch} - \upsilon^{loss} B_t^{Batt} \tag{9}$$

3.2.3 Solar Panels

The solar power (P_t^{Solar}) is limited by maximum installed power of solar panels (CP^{Solar}). Also, a curve (c_t^{Solar}) is multiplied in maximum power at each hour to provide the solar power amount.

$$P_t^{Solar} = c_t^{Solar} \times \mathrm{CP}^{Solar} \tag{10}$$

3.3 FDI Power

As mentioned in Sect. 2, there are two types of FDI, as follows:

3.3.1 Battery Depletion

The parameter P_t^{PD} enhances the amount of load as a result of FDI (11). This parameter cannot be larger than a percentage of load to not be detected and prevented. $\sigma_s^{\mathrm{PSFDI}}$ defined the attack severity in each scenario.

$$0 \leq P_t^{\mathrm{PD}} \leq \sigma_s^{\mathrm{PSFDI}} P_t^L \tag{11}$$

3.3.2 Load Shedding

Similar to power depletion, Eq. (12) defines the amount of load changes through load shedding. It is also limited by a special amount of load at each hour to be undetectable.

$$0 \leq P_t^{\mathrm{LS}} \leq \sigma_s^{\mathrm{PSFDI}} P_t^L \tag{12}$$

4 Simulation Results

In this section, the effects of FDI cyberattack on (1) battery depletion and (2) load shedding are considered in a typical day. Solar PV power is illustrated in Fig. 2, and the electrical load is shown in Fig. 3. The available charge of the battery is displayed in Fig. 4. It is assumed that there is an available charge in the battery in the first hours of the day, and the battery is charged by solar PV power during the day. The general algebraic modeling system (GAMS) software is utilized for simulation purposes.

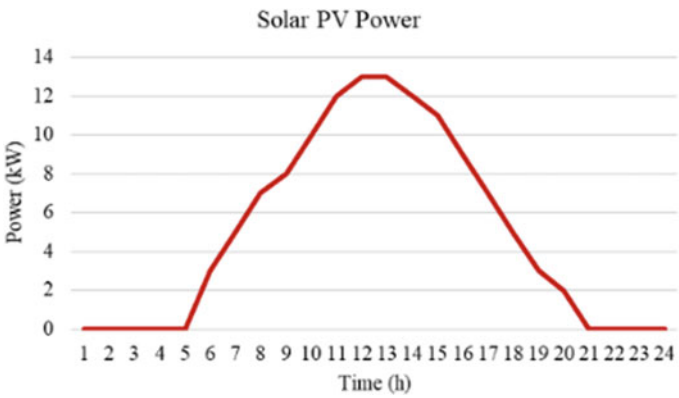


Fig. 2 Solar PV power in 24 h of a day

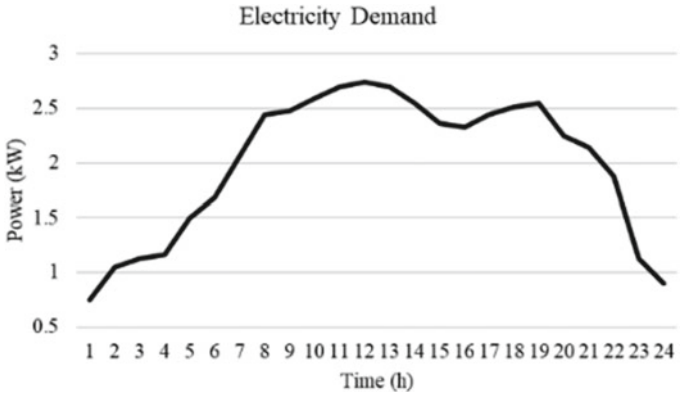


Fig. 3 Electricity demand of ISS

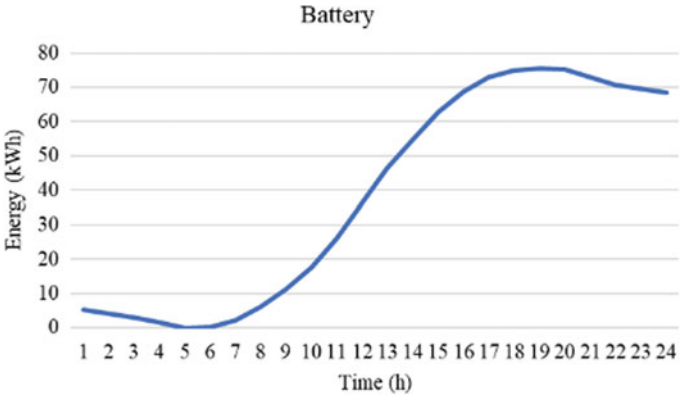


Fig. 4 Available charge/energy in the battery

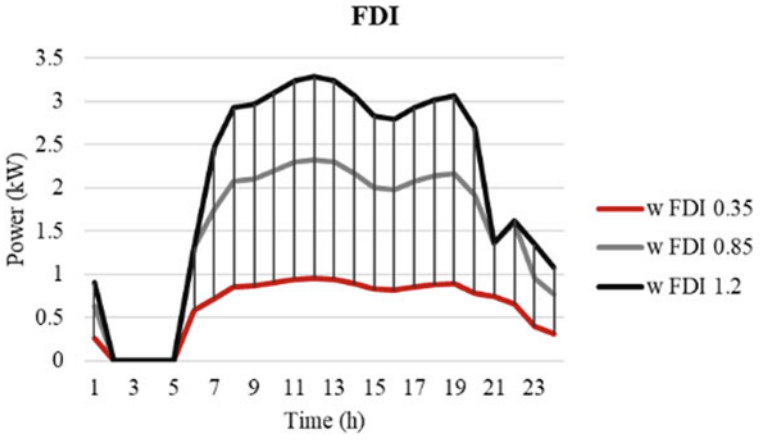


Fig. 5 FDI for the battery depletion

4.1 FDI on Battery Depletion

The impact of FDI on battery depletion of ISS power system: Fig. 5 shows the amount of FDI that has been injected into the electricity demand of ISS. Figure 7 illustrates how much battery is depleted by applying the FDI in Fig. 5. As it can be observed from Figs. 5 and 6, the electricity demand in Fig. 6 has been increased by applying the FDI in Fig. 5. The more FDI is injected into the ISS power system, the more electricity demand is increased. The coefficients of 0.35, 0.85, and 1.2 FDI in three scenarios have been injected into the system, and the electricity demand has increased. As a result, the battery depletes by applying the increase of FDI to the ISS power system.

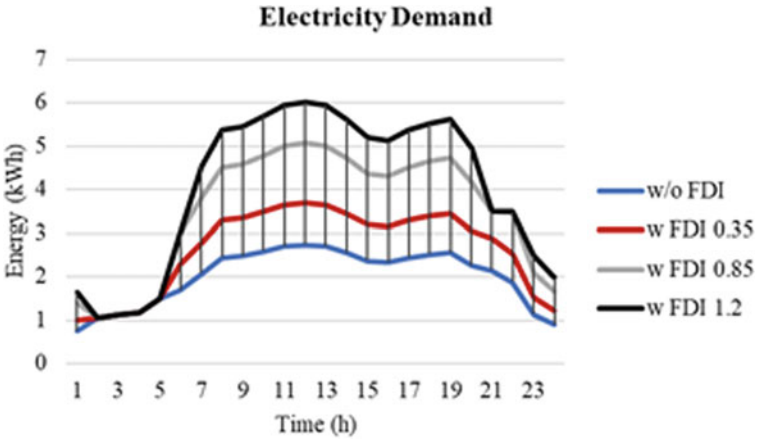


Fig. 6 Electricity demand under attacks

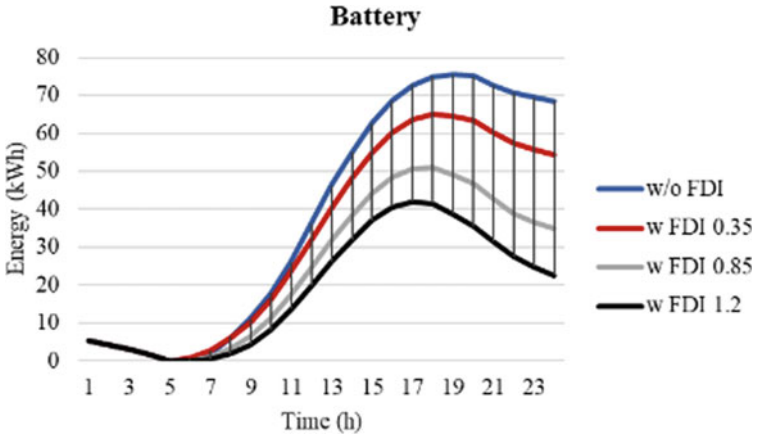


Fig. 7 Battery depletion under different cyberattacks in different scenarios

4.2 FDI on Load Shedding

The impact of FDI on load shedding of ISS power system: The amount of FDI injected into the power system is illustrated in Fig. 8, and the reduced electricity demand as the result of the FDI, shown in Fig. 8 and is displayed in Fig. 9. The coefficients of 0.1, 0.3, and 0.6 FDI have been injected into the system in three different scenarios. As it can be observed from Fig. 9, the electricity demand is reduced when FDI is injected into the ISS power system. The more FDI is injected into the system, the more electricity demand is reduced. However, there is some distortion in electricity demand shape (Fig. 8). If the FDI increases, it ascends the probability of being

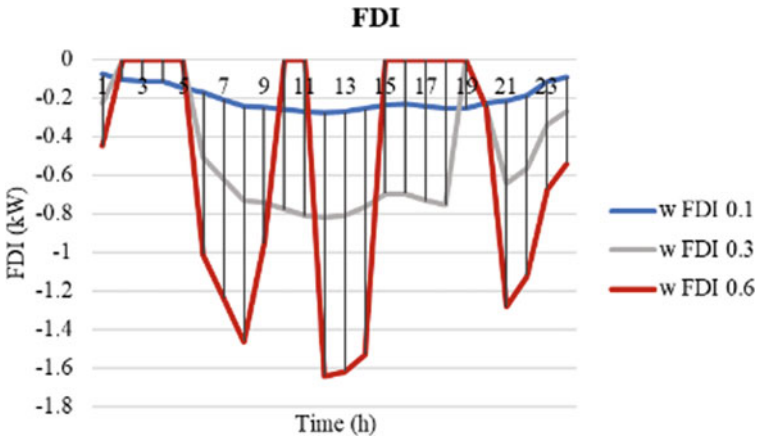


Fig. 8 FDI for load shedding attack

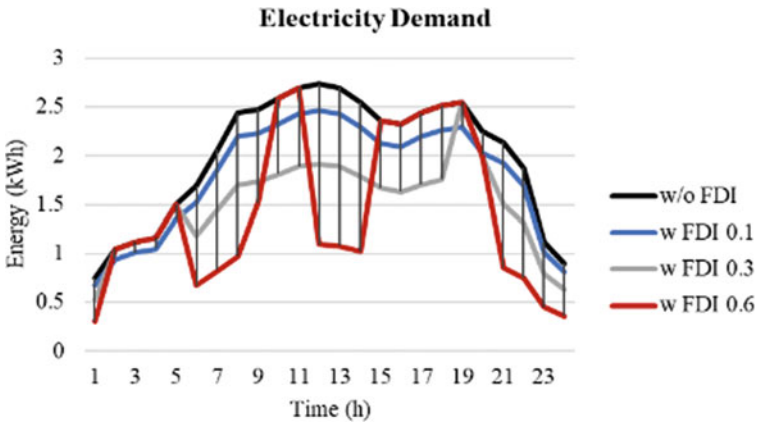


Fig. 9 Electricity demand under load shedding

detected. Hence, the attacker should consider the margin of FDI in order not to be detected.

5 Defense Mechanisms/Solutions

Considering the aforementioned vulnerabilities of the ISS, the attacker can breach/compromise the system through the Earth station and communication channel between the Earth station and space station. To protect the ISS against the cyber-attacks such as the FDI cyberattack on the ISS power system, two defense mechanisms/solutions, as shown in Fig. 10, are presented in this section.

5.1 Securing the Earth Station

The proposed mechanism for securing the Earth station against the attacks is to use strong mathematical model for cryptography, so it prevents the attacker to breach the confidential information of system (confidentially attack) and tamper the information (integrity attack). Furthermore, Firewall is recommended to monitor incoming traffic based on the organization's previously established security policies and keep dangerous traffic out. An intrusion prevention system (IPS) is suggested to identify malicious activities, report the threats, and prevent them to compromise the system.

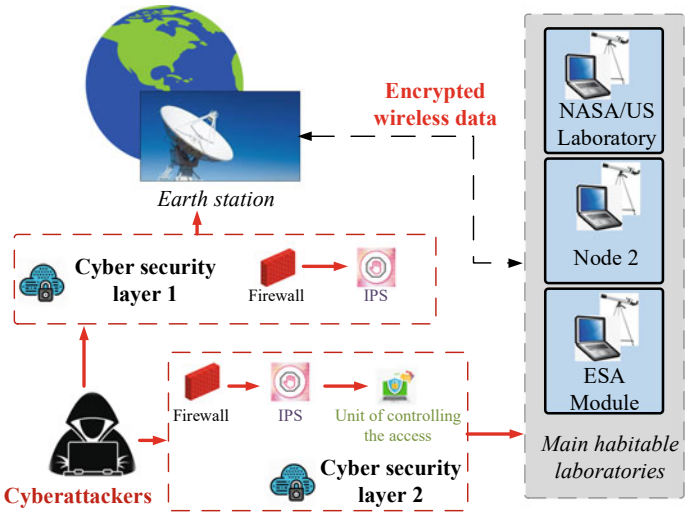


Fig. 10 Proposed cybersecurity layers for ISS against FDI

5.2 Securing the Communication Channel Between the Earth and Space Station

The proposed mechanism for communication channel between the Earth and space station against the attack is cryptography using strong mathematical model to prevent the confidentially and integrity attacks. Adding control access to the ISS management system prevents unauthorized person or unauthentic devices are connected to the system, so availability of the ISS will be enhanced. Furthermore, Firewall and IPS are recommended to be added to the ISS management system.

6 Conclusion

ISS, as a research center satellite in space, orbits the Earth with the purpose of testing spacecraft systems/equipment for long-term mission to Mars, Moon, etc. Power system of the ISS including battery and solar PV plays an important role in all satellites. The ISS is controlled through on-board algorithms or via commands transferred through the wireless channel from the Earth station. Vulnerabilities of the Earth station and wireless channel as well as globalizing the space industry/tourisms open doors for different cyber threats. Considering cybersecurity challenges of the ISS, we proposed two potential FDI cyberattacks resulting in (1) battery depletion and (2) load shedding. Scrutinizing the effect of the FDI on ISS, we proposed some defense mechanisms to prevent/protect the ISS against the cyberattacks.

References

1. http://www.spacex.com/
2. http://www.blueorigin.com/
3. http://www.virgingalactic.com/
4. http://www.nasa.gov/
5. Duan T, Dinavahi V (2021) Starlink space network-enhanced cyber–physical power system. IEEE Trans Smart Grid 12(4):3673–3675
6. Falco G (2020) When satellites attack: satellite-to-satellite cyber attack, defense and resilience. In: Aerospace Research Central, American Institute of Aeronautics and Astronautics
7. Manulis M, Bridges CP, Harrison R et al (2021) Cyber security in new space. Int J Inf Security 20:287–311
8. Vessels L, Heffner K, Johnson D (2019) Cybersecurity risk assessment for space systems. In: IEEE space computing conference (SCC), pp 11–19
9. Zheng W, Meng X, Jiang L (2008) Security issue for space internet. In: IEEE international conference on intelligent computation technology and automation (ICICTA), pp 194–198
10. Takamura E, Mangum K, Wasiak F, Gomez-Rosa C (2015) Information security considerations for protecting NASA mission operations centers (MOCs). In: IEEE aerospace conference, pp 1–14

11. Fischer D, Merri M, Engel T (2018) Security extensions for space-link communication. In: Proceedings of 17th IEEE international conference on computer communications and networks, pp 1–6
12. Mailloux LO, Span M, Mills RF, Young W (2019) A top-down approach for eliciting systems security requirements for a notional autonomous space system. In: IEEE international systems conference (SysCon), pp 1–7
13. Klesh AT, Cutler JW, Atkins EM (2012) Cyber-physical challenges for space systems. In: IEEE/ACM third international conference on cyber-physical systems, pp 45–52
14. Mailloux LO, Mills RF (2020) Autonomous space resupply vehicle systems security design principle case study. In: IEEE conference on communications and network security (CNS), pp 1–8
15. Komatsu M, Yanabu S (2008) Analysis of the small signal stability for the international space station/JEM electric power systems. In: IEEE 2nd international power and energy conference, pp 106–111
16. Fraas LM, Landis GA, Palisoc A, Jaffe P (2016) Space mirror development for solar electric power and the international space station. In: IEEE 43rd photovoltaic specialists conference (PVSC), pp 2558–2561
17. Schwanbeck E, Dalton P (2019) International space station lithium-ion batteries for primary electric power system. In: IEEE European space power conference (ESPC), pp 1–1

New Approach to Rural Energy Planning Based on ICHC-M_{GK}

Hery Frédéric Rakotomalala, Eric Jean Roy Sambatra, André Totohasina, and Jean Diatta

Abstract This paper aims to provide a tool to assist rural energy planning in developing countries. Indeed, classical methods seem to be limited for the implementation of power generation plants in these often remote and isolated areas. These localities are home to populations living on less than USD2 (2$) per day with a typical and complex conditions for the implementation of power plants. Thus, on the basis of surveys carried out in several villages in the northern region of Madagascar, a data classification method is used in order to provide electrification operators with a tool to help them make decisions on the potential technico-economic and cultural feasibility of electrification projects. The implicative and cohesive hierarchical classification method according to the interestingness measure M_{GK} (ICHC-M_{GK}), which has already proven itself in several didactic studies, is used herein for an economic-socio-energy study of households in some rural areas of the DIANA region of Antsiranana, in order to highlight the most used and consumed sources of energy in rural areas of the northern area of Madagascar. The results are conclusive because they highlight the potential of the sites in terms of electrification. Also, the approach makes it possible to make the link between the economic activities, the cultural practices of the villages, and the various elements relating to the energy component.

Keywords Rural electrification · Energy planning · Developing countries · Association rule · Cohesion · Clustering · Dendrogram

H. F. Rakotomalala (✉) · A. Totohasina
Department of Mathematics and Computer Science, University of Antsiranana, Antsiranana, Madagascar
e-mail: fredericrakotomalala@yahoo.fr

A. Totohasina
e-mail: andre.totohasina@gmail.com

E. J. R. Sambatra
Department of Electrical Engineering, Institute of Technology, Diego Suarez, Madagascar
e-mail: ericsambatra@gmail.com

J. Diatta
Department of Computer Science and Mathematics, University of La Réunion, Saint Denis, France
e-mail: jean.diatta@univ-reunion.fr

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_12

1 Introduction

According to World Bank report, 3.3 billion people are estimated to live in rural areas worldwide, just under half of the world's population. And in Africa, which accounts for about 15% of the world's population, the consumption of electrical energy is estimated at only 3% of the world's consumption. More than 600 millions peoples live without electricity in Africa, more than 80% of whom live in rural areas. Only two countries in the region, Mauritius and Seychelles, have achieved near-universal coverage. The household access rate to electricity is greater than or equal to 75% in only six African countries. Nearly two-thirds of countries in the region have access rates below 50% [1]. Moreover, there is a huge disparity in terms of electrification because while the rate for the continent as a whole is estimated at 42%, only 25% of rural areas are concerned. The fact is even more alarming in the case of sub-Saharan rural areas where the electrification rate does not exceed 10%.

The case of Madagascar is not at all isolated from the case of Africa and more particularly the sub-Saharan zone. The rate of electrification is very low as it is estimated at only 15% for the whole country. And among these 15% electrified and mainly urban areas, rural areas only benefit from 6%. Therefore, in a more direct way, only 0.9% of rural areas in Madagascar benefit from electricity produced by operators recognized by the state. Although the energy policy developed by the Ministry of Energy has taken into consideration these areas through the establishment of state organizations such as the Agency for the Development of Rural Electrification (ADER), their achievements are very limited, and the current programs would not reverse this situation which certainly contributes to the economic development of the country. This policy gap is reinforced by other factors such as the socio-cultural and economic situations of the villages as well as their geographical distribution [2]. Indeed, for a population of 25 millions inhabitants occupying a very vast area of 587,000 km^2, the population density per km^2 is very low. It amounts to only less than 34 inhabitants/km^2. This is therefore one of the problems to be taken at face value in the case of rural electrification in southern countries because villages are scattered. Another blocking factor is that the annual income of the inhabitants does not allow them to contribute to any direct investment for an electrification project. The estimate suggests an annual per capita income of less than USD 450. The barriers are multiple and are found at all levels, but in this paper, we rather address the basic technical component that could influence the energy policy of the country. Indeed, even if efforts are made by the structures in place for the implementation of strategic planning, the above-mentioned elements limit the effectiveness of this planning.

This article thus aims to provide decision and development actors with a tool to help them make more efficient energy planning through a classification approach with a statistical method called ICHC-M_{GK} [3]. The idea is to start from a multidimensional database of socio-economic and cultural studies on several villages not benefiting from electrification, and to link them with the potential electrification and especially the most influential variables in order to better guide the planning and development strategies to develop the energy sector.

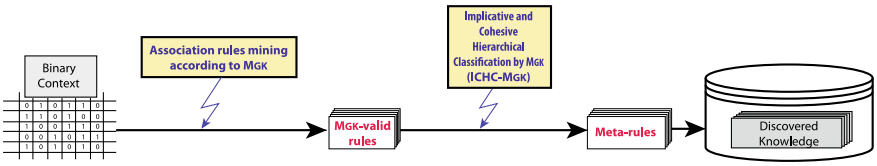


Fig. 1 Methods for data mining process according to M_{GK}

The remaining of the paper is structured in three main sections. We present the first, the ICHC-M_{GK} concept, then its application to survey data from about several villages in the Diana region of northern Madagascar, and finally, a section is reserved for the analysis of the results and especially its contribution to the energy component.

2 Materials and Method

Association rules mining [4] is one of the most popular data mining techniques. Association rules are useful for discovering relationships within very large databases. A new algorithm for extracting association rules M_{GK}-valid was born [5] based on the works of [6] as well as the unifying theory of normalized measures and normalizable ones, case of M_{GK}, also a new ICHC-M_{GK}-tool by referring to the theory of statistical implicative analysis—(SIA) in [7] and implicative and cohesive hierarchical classification in [8], whose adopted methods are detailed in [3, 9] (Fig. 1).

2.1 Data Collection

Surveys around 1736 households, on the type of their habitats, their economic activities, the sources of energy, and the electrical equipment, they use were carried out in three rural areas of Antsiranana, including AREA1 (Ambilobe and its surroundings), AREA2 (Peri-Urban of Diego Suarez), and AREA3 (Ambanja and its surroundings) (Cf. Table 1). The collected paper versions were then digitized on Excel to have a binary database (Cf. Table 2). Finally, the binary database obtained in .csv format was processed by the new ICHC-M_{GK}-tool for analysis. The data are composed of 120 variables were digitized and transformed into binaries in Excel (Cf. Table 2).

2.2 Used Methods for Data Mining

Let $(\mathbb{K}, \mathcal{H}, \mathcal{I})$ be a binary context. Consider a set of households $\mathcal{H} = \{h_1, h_2, \ldots, h_{1736}\}$ and a set of items $\mathcal{I} = \{i_1, i_2, \ldots, i_{120}\}$ (Cf. Table 2). The association rule between

Table 1 Survey sheet by area

			AREA1	AREA2	AREA3
TYPICAL HOUSE		Hard	+	+	+
		Sheet	+	+	+
		Falafa	+	+	+
		Bamboo	+	+	+
		Timber	+	+	+
ECONOMIC ACTIVITY	AGRICULTURE	Rice	+	+	+
		Pistachio	−	+	−
		Tomato	+	−	−
		Vanilla	−	−	+
		Cocoa	+	−	+
		Coffee	−	−	+
		Sugar cane	+	−	−
		Pepper	−	−	+
		Cassava	+	+	−
		Vegetable	−	+	−
		Banana	+	−	+
		Orange	−	−	+
		Mango	−	−	+
		Bean	+	−	+
		Cotton	+	−	−
		Corn	−	+	−
	FISHING	Shrimp	−	−	+
		Fish	−	−	+
	FARMING	Pig	−	−	+
		Ewe	−	−	+
		Beef	+	+	+
		Poultry	+	+	+
	COMPANY	Grocery store	+	+	+
		Bar	+	−	+
		Resto	+	−	+
		Video	+	−	+
		Handicraft	+	+	+
		Tourism	−	−	−

		AREA1	AREA2	AREA3
SOURCE OF ENERGY USED	Candles	+	−	+
	Oil	+	+	+
	Dry wood	+	+	+
	Coal	+	+	+
	Gas	+	+	+
	Accumulators	+	+	+
	Battery	+	+	+
	Solar panel	+	+	+
	Gasoline group	+	+	+
	Diesel group	−	−	−
EXISTING ELECTRICAL EQUIPMENT	Lamp	+	+	+
	T.V	+	+	+
	Radio	+	+	+
	Iron	+	+	+
	Player	+	+	+
	Telephone	+	+	+
	Fridge	+	+	+
	Mill	−	−	+

Table 2 Extract of binary table 120×1736

	AREA1	…	Fridge_no	Mill_yes	Mill_no	Poor	Average	Rich
1	1	…	1	0	1	1	0	0
2	1	…	1	0	1	1	0	0
3	1	…	1	0	1	1	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1734	0	…	1	0	1	0	0	1
1735	0	…	1	0	1	0	0	1
1736	0	…	1	0	1	0	0	1

two Boolean variables is defined by the analysis of the contingency table constructed by crossing them (Cf. Table 3).

Mathematical modeling: consider a finite discrete probability space $(\mathcal{H},\mathcal{P}(\mathcal{H}),P)$ such that for any event X of $\mathcal{P}(\mathcal{H})$, $P(X) = \frac{\text{card}(X)}{\text{card}(\mathcal{H})}$. Let us note the set of n households, on which we have measured m random variables of Bernoulli, and let $\mathcal{I}$ the set of items $\mathcal{I} = \{i_1, i_2, i_3, \ldots, i_n\}$. For all $X \in \mathcal{P}(\mathcal{I}) \setminus \{\varnothing; \mathcal{I}\}$, for all $x_i \in X$, x_i is an application from $\mathcal{H}$ to $\{0; 1\}$ and $P(x_i = 1) = \frac{(\text{card}(x_i^{-1}(1)))}{n}$, with $n = \text{card}(\mathcal{H})$. Any non-empty part of $\mathcal{I}$ will be called an $\mathcal{I}$ pattern. Then for A pattern, $A' = A^{-1}(1)$ and $n_A = \text{card}(A')$. For A and B patterns, $n_{AB} = \text{card}(A' \cap B')$ the number of transactions that realize both A and B. Let $\overline{A} = \mathcal{I} - A$ be the logical negation of a pattern

Table 3 Contingency and associated joint probability table

	B	$\overline{B}$	$\sum$
A	n_{AB}	$n_{A\overline{B}}$	n_A
$\overline{A}$	$n_{\overline{A}B}$	$n_{\overline{A}\,\overline{B}}$	$n_{\overline{A}}$
$\sum$	n_B	$n_{\overline{B}}$	n

$\Longrightarrow$

	B'	$\overline{B'}$	$\sum$
A'	$P(A' \cap B')$	$P(A' \cap \overline{B'})$	$P(A')$
$\overline{A'}$	$P(\overline{A'} \cap B')$	$P(\overline{A'} \cap \overline{B'})$	$P(\overline{A'})$
$\sum$	$P(B')$	$P(\overline{B'})$	$\frac{\text{Card}(\mathcal{H})}{\mathbf{n}} = 1$

A. The real number $P(A')$ will be called the *support* of the pattern A denoted $\text{supp}(A) = \frac{\text{card}(A')}{n}$ [4].

An interestingness probabilistic measure is a real function μ of $\mathcal{P}(\mathcal{I}) \times \mathcal{P}(\mathcal{I})$ such that for any association rule $A \to B$, the value of $A \cap B = \oslash$, $\mu(A \to B)$ is calculated from the four quantities $n = \text{card}(\mathcal{H})$, $P(A')$, $P(B')$, and $P(A' \cap B') = \text{supp}(A \cup B)$. Finally, for two patterns (or items) A and B of a binary context, the interestingness measure M_{GK} is defined by

$$M_{GK}(A \to B) = \begin{cases} M^f_{GK}(A \to B) = \frac{P(B'|A') - P(B')}{1 - P(B')}, & \text{if } A \text{ favors } B \text{ (favoring component);} \\ 0, & \text{if } A \text{ and } B \text{ are independant;} \\ M^d_{GK}(A \to B) = \frac{P(B'|A') - P(B')}{P(B')}, & \text{if } A \text{ disfavors } B \text{ (defavoring component)} \end{cases}$$

The theoretical work published in [3, 9] allowed us to develop an algorithm for extracting M_{GK}-valid association rules [5].

- The extraction of the association rules is based on the interestingness measure M_{GK} compared to the relation two by two of the variables according to contingency tables (Cf. Table 4); and the validation of the extracted rules is done in relation to the favoring component M^f_{GK} which is implicative and the critical value $M^f_{GK(\alpha)}$ having a relation with the χ^2 of degree of freedom 1 at the risk threshold α chosen by ourselves, such as $M^f_{GK} > M^f_{GK(\alpha)}$ with $M^f_{GK(\alpha)} = \sqrt{\frac{1}{n}\frac{n-n_A}{n_B}\frac{n_B}{n-n_B}\chi^2_{\text{Theoritical}(\alpha)}}$. In our case, $\alpha = 10\%$, ones ponding to $\chi^2 = 2.7$;
- The value of support according to M^f_{GK} such as $\text{supp}_{M^f_{GK}}(A \to B) = \text{supp}(A)\left[(1 - \text{supp}(B))M^f_{GK}(A \to B) + \text{supp}(B)\right]$ is generally seems low [9]. It is therefore essential to normalize this value to contrast it [5], and we denote it $\text{supp}^f_{(n)M_{GK}}$ with

$$\text{supp}^f_{(n)M_{GK}}(A \to B) = \frac{\text{supp}_{M^f_{GK}}(A \to B) - P(A')P(B')}{P(A')(1 - P(B'))}.$$

- The value of $\text{supp}^f_{(n)M_{GK}} \in \left]0.5, 1\right]$ allows us to establish the value of the cohesion between two items, denoted $\text{coh}_{\text{supp}_{(n)M_{GK}}}$ [5] with

Table 4 Extracted from contingency tables

	Handicraft_non	non_Handicraft_non
nb_fam__less5	905	3
non_nb_fam__less5	827	1
*		
	Candles	non_Candles
nb_fam__less5	895	13
non_nb_fam__less5	814	14
*		

$$\text{coh}_{\text{supp}_{M_{\text{GK}}}}(A, B) = \begin{cases} \sqrt{1 - (\text{supp}^f_{(n)M_{\text{GK}}}(A \rightarrow B))^2} & \text{, if } \text{supp}^f_{(n)M_{\text{GK}}}(A \rightarrow B) > 0,5 \\ 0 & \text{, if } \text{supp}^f_{(n)M_{\text{GK}}}(A \rightarrow B) \leq 0,5 \\ 1 & \text{, if } \text{supp}^f_{(n)M_{\text{GK}}}(A \rightarrow B) = 1 \end{cases}$$

- The implicative and cohesive hierarchical classification method according to the interestingness measure M_{GK} (ICHC-M_{GK}) is based on cohesion $\text{coh}_{\text{supp}_{(n)M_{\text{GK}}}}$ [3].

So we now process the data transformed into binaries of Table 2 by the ICHC-M_{GK} software for analysis.

3 Results and Interpretation

3.1 Results Obtained by ICHC-M_{GK}

From our concrete binary context with 1736 households and 120 variables, concerning rural electrification in 3 regions, we obtain 7140 valid association rules according to M_{GK}, with a risk threshold of 10% set by ourselves, of which, 3449 are positive rules and 3691 negative rules, with $0.0000003 \leq \text{supp}^f_{(n)M_{\text{GK}}} \leq 1$ which form 280 oriented pairs (OP), that is to say, $\text{card}(\text{coh}_{\text{supp}_{(n)M_{\text{GK}}}}) = 280$ and $0.01 \leq \text{coh}_{\text{supp}_{(n)M_{\text{GK}}}} \leq 1$.

Table 5 shows the values of inter-classes cohesions $\text{coh}_{\text{supp}_{(n)M_{\text{GK}}}}(\underline{A})$, inter-class implications $\psi(\underline{A}, \underline{B})$ where $\underline{A}$ forms the premise class of the meta-rule R on level of hierarchy $H(R(H))$ and $\underline{B}$ its antecedent class, the significant levels $l(\Omega, H)$, and the significant nodes $n(\Omega, H)$ of each level of the hierarchy H (Cf. paper [3]). We also present in Table 6 the 44 meta-rules $R(H)$ at the risk threshold $\alpha = 10\%$ and the extract of dendrogram representing the so meta-rules (Cf. Fig. 2).

3.2 Some Interpretations

Culture, animal husbandry, and fishing are the main activities of the rural inhabitants. Each selected and surveyed area have its specificity in relation to these three activi-

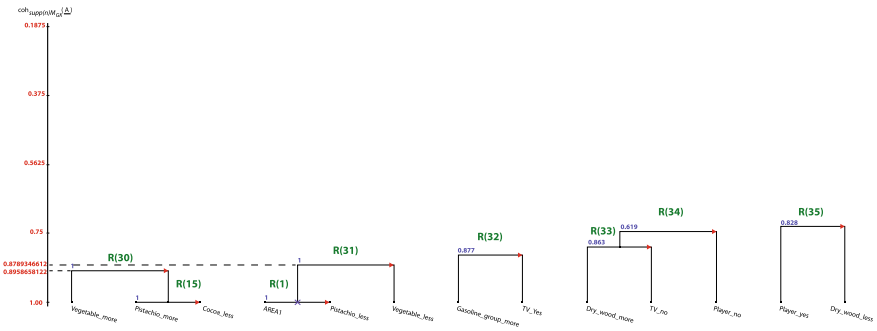


Fig. 2 Extract from the dendrogram representing the obtained meta-rules

Table 5 Table showing of each classes level, $\text{coh}_{\text{supp}_{(n)M_{GK}}}(\underline{A})$, $\psi(\underline{A}, \underline{B})$, significant level $l(\Omega, H)$, and significant node $n(\Omega, H)$

Level H	Inter-class cohesion $\text{coh}_{\text{supp}_{(n)M_{GK}}}(\underline{A})$	Inter-classe implication $\psi(\underline{A}, \underline{B})$	Significant level $l(\Omega, H)$	Significant node $n(\Omega, H)$
1	1.0000000000	1.0000000000	0.0001212921	0.0001212921
2	1.0000000000	1.0000000000	0.0000000085	−0.0001212836
3	1.0000000000	1.0000000000	−0.0000404137	−0.0000404222
...	...	...	...	...
30	0.8958658122	1.0000000000	0.0003085193	0.0006902468
31	0.8789346612	1.0000000000	0.0004128346	0.0001043153
...	...	...	...	...
42	0.1270000000	0.5533770243	−0.0014257252	0.0004735949
43	0.0650000000	0.5265454545	−0.0011513844	0.0002743409
44	0.0620000000	0.5258126195	−0.0009724098	0.0001789746

ties (Cf. Table 1). AREA1, for example, focuses on the cultivation of tomatoes, sugar cane, and cotton; AREA2 on the cultivation of pistachios, vegetables, and corn, while AREA3 on the cultivation of vanilla, coffee, pepper, oranges, and mangoes. This observation is supported by the rules $R(1)$, $R(2)$, $R(3)$, $R(30) \equiv$ Vegetable_more $\Rightarrow$ $R(15)$ and $R(31) \equiv R(1) \Rightarrow$ Vegetable_less of Table 6, and each class forming these respective rules have strong cohesion $\text{coh}_{\text{supp}_{(n)M_{GK}}}[R(1), R(2), R(3)] = 1$, $\text{coh}_{\text{supp}_{(n)M_{GK}}}[R(30)] = 0.8958658122$ and $\text{coh}_{\text{supp}_{(n)M_{GK}}}[R(31)] = 0.8789346612$ (Cf. Table 5):

- `R(1) = (AREA1 ⇒ Pistachio_less)`
- `R(2) = (Pistachio ⇒ AREA2)`
- `R(3) = (AREA3 ⇒ Tomato_less)`
- `R(30) = (Vegetable_more ⇒ (Pistachio_more ⇒ Cocoa_less))`
- `R(31) = ((AREA1 ⇒ Pistachio_less) ⇒ Vegetable_less)`

Table 6 The 44 meta-rules after clustering

Meta-rules	
R(1) = (AREA1 ⇒ Pistachio_less)	R(23) = (Battery_yes ⇒ Banana_less)
R(2) = (Pistachio ⇒ AREA2)	R(24) = (Cotton ⇒ Radio_no)
R(3) = (AREA3 ⇒ Tomato_less)	R(25) = (Resto_no ⇒ Shrimp_less)
R(4) = (nb_fam__less5 ⇒ Sugarcane_less)	R(26) = (Shrimp_more ⇒ Resto_yes)
R(5) = (Tomato_more ⇒ nb_fam__5_a_10)	R(27) = (Oil ⇒ Ewe_less)
R(6) = (Room_less3 ⇒ Orange_less)	R(28) = (Beef_less ⇒ Mill_no)
R(7) = (Orange_more ⇒ Room_3_a_4)	R(29) = (Mill_yes ⇒ Beef)
R(8) = (Orange ⇒ Room_more4)	R(30) = (Vegetable_more ⇒ (Pistachio_more ⇒ Cocoa_less))
R(9) = (Vanilla_more ⇒ House_Sheet)	R(31) = ((AREA1 ⇒ Pistachio_less) ⇒ Vegetable_less)
R(10) = (House_Falafa ⇒ Corn_less)	R(32) = (Gasoline_group_more ⇒ TV_yes)
R(11) = (House_Timber ⇒ Poor)	R(33) = (Dry_wood_more ⇒ TV_no)
R(12) = (Rice_less ⇒ Mango_less)	R(34) = ((Dry_wood_more ⇒ TV_no) ⇒ Player_no)
R(13) = (Pig_more ⇒ Rice_more)	R(35) = (Player_yes ⇒ Dry_wood)
R(14) = ((Pig_more ⇒ Rice_more) ⇒ Rich)	R(36) = (Sugarcane ⇒ House_Hard)
R(15) = (Pistachio_more ⇒ Cocoa_less)	R(37) = (Solar_panel_yes ⇒ Lamp_yes)
R(16) = (Tomato ⇒ Accumulator_more)	R(38) = (Lamp_no ⇒ Solar_panel_no)
R(17) = (Cocoa ⇒ Battery_no)	R(39) = (Sugarcane_more ⇒ nb_fam__more10)
R(18) = (Ewe_more ⇒ Cocoa_more)	R(40) = (Fish_more ⇒ Oil_more)
R(19) = (Pepper ⇒ Telephone_yes)	R(41) = (Gasoline_group_less ⇒ Bar_no)
R(20) = (Pepper_more ⇒ Radio_yes)	R(42) = (Rice ⇒ Average)
R(21) = (Coal ⇒ Cassava_less)	R(43) = (Poultry_less ⇒ Bean_less)
R(22) = (Cassava_more ⇒ Coal_more)	R(44) = (Telephone_no ⇒ Pepper_less)

For fishing, only AREA3 intensively practices this activity. Likewise for breeding, AREA3 practices ewe, pig farming, and it is the only area that has mill-manufactory in his riding. Household lifestyles depend on the most practiced economic activities in each area and vice versa. Thus, we will interpret the following rules:

- R(4) = (nb_fam__less5 ⇒ Sugarcane_less)
- R(5) = (Tomato_more ⇒ nb_fam__5_a_10)
- R(6) = (Room_less3 ⇒ Orange_less)
- R(7) = (Orange_more ⇒ Room_3_a_4)
- R(8) = (Orange ⇒ Room_more4)
- R(9) = (Vanilla_more ⇒ House_Sheet)
- R(10) = (House_Falafa ⇒ Corn_less)
- R(11) = (House_Timber ⇒ Poor)
- R(36) = (Sugarcane ⇒ House_Hard)

Households less than 5 people do not practice cultivation of sugar cane (rule $R(4)$). In other words, AREA2 and AREA3 are made up of households of less than 5 people. While in AREA1, households are made up of more than 5 people, because the culture of sugar cane needs manpower, and they live in hard houses $R(36)$. The rule $R(5)$: those who grow more tomatoes (the AREA1) are families of more than 5–10 people and live generally in Falafa houses.

The households in AREA3 who grow vanilla and oranges live in sheet metal houses and more than 3 rooms (rules $R(6)$, $R(7)$, $R(8)$ and $R(9)$). While the habitats in AREA1 are built in Falafa and less than 3 rooms (Cf. Table 1, $R(10)$). We observed as well as the families who live in Timber houses are classified as poor families $R(11)$ = (House_Timber ⇒ Poor).

Generally, households that raise more pigs grow a lot of rice and are part of rich families $R(14) \equiv (R(13) \Rightarrow \text{Rich})$.

The rules $R(16)$, $R(17)$, $R(21)$, $R(22)$, $R(23)$, $R(27)$, $R(32)$, $R(33)$, $R(34)$, $R(37)$, $R(38)$, $R(40)$, $R(41)$ reflect the use of household energies in relation to their activities. Households growing tomatoes (AREA1) generally use batteries as a source of energy (rule $R(16)$). Households (AREA1, AREA2) that use more charcoal as a source of energy are cultivators of cassava (rule $R(22)$). Households (AREA2) who cultivate the banana crop generally use batteries (rule $R(23)$).

The areas (AREA1, AREA3) that have more Bar consume more gasoline, and households generally have TV as household electrical equipment (rules $R(31)$, $R(42)$). The fishermen's areas (AREA3) consume more oil (rule $R(40)$). Solar panels are generally used for lighting households or villages (rules $R(37)$, $R(38)$). Households that consume more dry wood do not have a TV but a reader as an electrical appliance (rules $R(33)$, $R(34)$).

Rice cultivation is therefore the indicator of the standard of living of rural households, since middle-income households cultivate rice sufficient for their family $R(42)$ = (Rice ⇒ Average).

We can then reclassify the 3 areas studied according to their most consumed energy sources.

Table 7 Classification of areas according to the energy sources used

	AREA1	AREA2	AREA3
Accumulator	+	−	−
Coal	+	+	−
Battery	−	+	−
Oil	−	−	+
Gasoline group	+	−	+

Table 7 shows that in addition to the consumption of dry wood, areas 1 and 2 use more coal as an energy source, which causes nuisance in the forests and would have a very serious environmental impact.

The approach developed in this article seems relevant and surprising because it allows the identification of hidden or hardly detectable information in the case of the use of classical approaches. As an example, we would like to specify here the link between the use of batteries and the cultivation of bananas. This information is crucial in the sense that the northern region of the country is quite fertile for the cultivation of tropical fruits, especially bananas, which are omnipresent throughout the year. Thus, in order to increase the purchasing power of households, they can invest in the massive planting of banana trees, which will be an additional income-generating activity for them. Obviously, this aspect will have its impact on human development such as access to services such as electricity.

4 Conclusion

The aim of this paper is to provide actors in the field of rural electrification, mainly in sub-Saharan Africa, with a tool to help optimizing the implementation of energy planning. The case of the Northern zone of Madagascar is the subject of experimentation. Thus, through a database collected on several representative communes, the ICHC-M_{GK} method is used. The approach experimented prior to other applications has shown its effectiveness in the treated field. The meta-rules established by the classification highlight, in a precise manner, the links between the different parameter of study, such as the professional activities of households, their locality and others, with energy-related characteristics. This could more effectively guide planners and even decision-makers in terms of socio-economic and energetic policy, because the rules highlight the economic activities to be promoted in the localities increasing their payment capacities in case of implementation of power plants production.

Finally, the proposed approach shows all its importance in the sense that it makes it possible to detect relevant or even surprising information that is difficult to identify using conventional methods. In terms of perspectives, we are considering extending our applications on a national and international level.

Acknowledgements This paper is realized due to the support of Ambassy of France in Madagascar and AUF.

References

1. Blimpo M, Cosgrove-Davies M (2020) Electricity access in Sub-Saharan Africa: uptake, reliability and complementary economic impact factors. https://openknowledge.worldbank.org/handle/10986/31333
2. Sambatra E, Andrianajaina T, Razafimahefa T, Heraud N (2016) An original optimization approach of the management of isolated sites power plants: case of the Ivovona Village—Madagascar. https://site.ieee.org/repc-2016/paper-titles-authors/abstract-37/
3. Rakotomalala HF, Ralahady B, Totohasina A (2018) A novel cohesitive implicative classification based on M_{GK} and application on diagnostic on informatics literacy of students of higher education in Madagascar. In: Third international congress on information and communication technology, ICICT, vol 797, pp 161–174
4. Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases. In: ACM SIGMOD International conference on management of data, vol 22, pp 207–216
5. Rakotomalala HF, Totohasina A (2018) An efficient new cohesion indice based on the quality measure of association rules M_{GK}. In: 2nd World conference on smart trends in system, security & sustainability
6. Feno D (2007) Mesures de qualité des règles d'association: normalisation et caractérisation des bases. Université de la Réunion. https://tel.archives-ouvertes.fr/tel-00462506
7. Gras R (2013) L'analyse statistique implicative: Méthode exploratoire et confirmatoire à la recherche de causalités. Cépaduès, 2ème édition
8. Ratsimba-Rajohn H (1992) Contribution à létude de la hiérarchie implicative: application à l'analyse de la gestion didactique des phénomènes d'ostension et de contradictions. U.F.R. de Mathématiques, Université de Rennes I
9. Rakotomalala HF, Totohasina A, Diatta J (2017) Extraction des règles d'associations M_{GK}-valides avec contribution de Support. In: Actes Des 24èmes Rencontres De La Société Francophone De Classification, pp 29–32

Texture Analysis and Feature Extraction in Tumor Skin Cancer: Survey

Asmaa Abdul-Razzaq Al-qaisi and Luay Edwar

Abstract Texture is a term used to characterize the surface of objects and region and represents main features in pattern recognition and image processing. The concept of image or object, and is defined as a function of the brightness's spatial fluctuation intensity of pixel or shape, analysis of texture may be classified into four distinct categories (structural, statistical, transform, and a model-based approach). The purpose of feature extraction is to convert an image to a matrix vector and to create a unique representation of signal values. The term "feature extraction" refers to the process of extracting features from an image without the use of any processing procedure. This article presents the main texture strategies extracted such as co-occurrence matrix, gradient, contrast, DCT, DWT, fractal, and PCA are used for analysis image skin tumors and compare, combine these strategies to reach a high diagnostic accuracy by computational complexity to reduce the challenge example (rotation, noise, etc.) to become familiar with the many sorts of features that may be utilized in DIA (digital image analysis) for future researchers are provided.

Keywords Image processing · Feature extraction · Texture analysis · Skin cancer

1 Introduction

For past decades, increasing skin cancer incidence rates in countries. For example, in UK, melanoma incidence has risen 119% in (1990) and in United State of America [1, 2] from (27,600) cases in 1990 to (91,270) cases through 2018, while more than 87,000 melanoma cases every year in United State [3]. Moreover, report nearly (13,000) cancer case in UK and Australian. Increases of infection rates are due to

A. A.-R. Al-qaisi (✉)
Informatics Institute for Postgraduate Studies, Baghdad, Iraq
e-mail: asma-72@coeduw.uobaghdad.edu.iq

L. Edwar
University of Information and Communication Technology, Baghdad, Iraq
e-mail: loayedwar57@uoitc.edu.iq

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_13

Table 1 The 7 points checklist [7]

Original (7PCL) (score ≥ suggest referral)	Weighted (7PCL) (score ≥ 3 suggest referral)
All feature (weight is equal)	**Main features (2 point)**
Pigmentation is irregular	Pigmentation is irregular
Change size of tumor	Change size of tumor
Inflammation	Irregular border
Border is irregular	**Minor features (1 point)**
Altered sensation or itch	Inflammation
Larger than other tumor (diameter larger than 7 mm)	Larger than other tumor (diameter larger than 7 mm)
Crusting of tumor	Itch or altered sensation
	Crusting of tumor

several reasons such as low of the ozone layer and direct exposure to sun light. Therefore, new technologies must be used for diagnosis skin cancer early.

There are two ways to acquire a color medical picture of a skin tumor, which is either the microscopic picture or the dermatoscopy, and the endoscopic picture correctly identifies the lesion. Macroscopic image is obtained using mobile phone or cameras [4]. Many methods are used for diagnosis the dermoscopic image such as (i) hierarchical process, this method focuses on distinction between non-melanocytic and melanocytic lesions, and subsequently, between benign and malignant lesion depends on series rules of dermoscopic criteria that proposed to avoid subjectivity analysis [5]; (ii) other methods focus on possible density and criteria inside the tumor pattern analysis [6], but other methods used criteria associated with melanoma, example (7 points checklist) original and weighted [7], and Menzies method [8]. The **(7PCL)** includes the following Table 1.

In **Menzies method,** it is a simplified dermatoscopy method for diagnosing melanomas [9]; in Menzies approach, this method consists of determining six colors of the skin image; tumors are characterized as either negative or positive. A tumor is called malignant if it has negative features (single color and pigmentation asymmetry). The benign tumors have sequential positive features such as (veil of blue and white as scar, radial syncope of false legs, black dots/balls in the periphery, multiple five or six colors, multiple gray dots, and multiple network), shown in Fig. 1. The Menzies method gives specificity (71%) and sensitivity (92%) for diagnosing melanoma [9].

As for the third method, it combines the rules of (ABCD) that includes the degree of (asymmetry, border structure, lesions, and color distribution [10]), and (CASH) includes (color, architecture, symmetry, homogeneity) [11]. Taking into account the melanoma and non-melanocytic cells in addition to (family history, age, gender), Fig. 2 shows the ABCD rule for diagnosing moles (healthy and unhealthy) [12, 13].

The research community trying to develop computer-aided design (CAD) system to be used by dermatologists [14, 15], through the use of image processing to identify

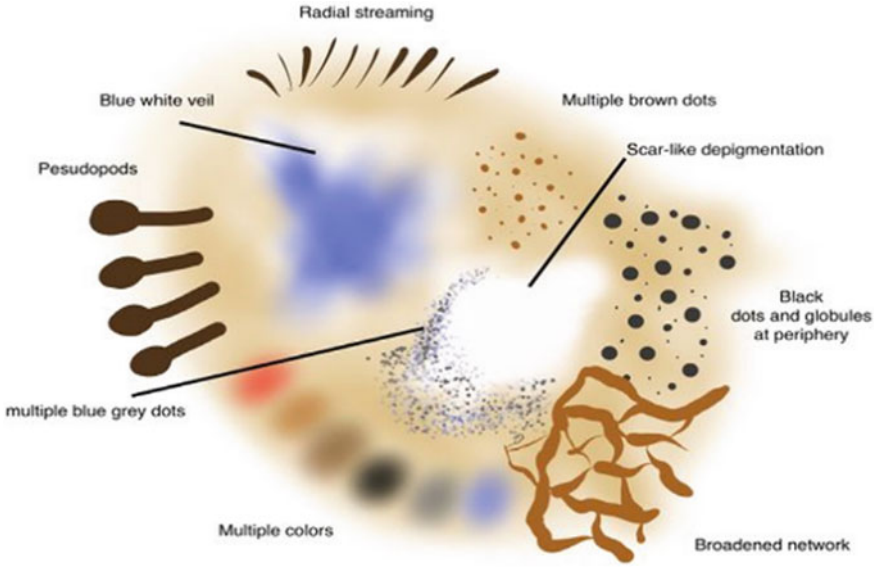


Fig. 1 Illustration of positive feature

Fig. 2 ABCD rule for nevi diagnosing (healthy and unhealthy)

and classify the type of the tumors. A series of steps that are carried out on medical images to reach an accurate diagnosis of the type of tumor, which are as follows (as shown in the Fig. 3).

Fig. 3 Series of steps those involved to perform medical image classification

Preprocessing is used to deal with images that do not have high quality for analysis purposes. The preprocessing is divided into two main steps (the first is image optimization, and the second is noise removal).

Image Enhancement It includes (calibration/normalization, contrast enhancement, color space conversion), and either software normalization (using image properties to perform normalization) or hardware includes (camera, color acquisition, and aperture) [16].

Contrast Enhancement Color space-related defects (not uniform and not device independent, so an alternative color space such as CIE L^*u^*v, CIE L^*a^*b, HSV and galleries are used.

Artifact Removal two types of artifact removal that is objects cutaneous in nature, such as (color texture, skin lines, hair, blood vessels) and acquisition artifacts, for example (air bubbles, reflection, black frames, ruler and ink markings), in this case, median filter is used [17, 18].

Lesion Segmentation is a technique for dividing an image into sections that have a high association to things or areas of the actual world. The variety of lesion (in size, shapes, colors) and different type of skins (white, brown, corny) and textures that make difficult to develop robust algorithm of segmentation to achieve proper and accurate extraction of features [19].

Feature Extraction determining the features either by hand-crafted features includes (shape, texture, symmetry, color) or by dictionary base feature, sparse coding, or bag of feature used to acquisition local description of tumor.

Feature selection this method is used to reduce the dimensionality of feature by eliminating redundant, irrelevant, noise features.

Lesion Classification Finally, algorithm training that used to predict a diagnosis such as ANN, SVM or other classifier.

2 Techniques for Feature Extraction

In medical image analysis, texture is important term used to define concept or objects of image, many approaches used to describe texture images accurately. This paper reviews various method to analyze the texture and extract the feature in details such as wavelet transform feature, gradient feature, direct cosine transform (DCT), contrast feature, co-occurrence feature, run length feature, principal component analysis-derived (PCA) features, and many techniques can be used to extract the feature from image explain as follows:

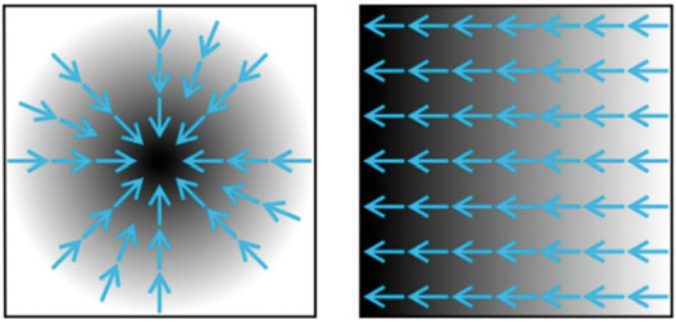


Fig. 4 Gradient direction density

Gradient Technique

There are two types of gradients, and the direction of the arrows is used to represent the gradient's direction, whereas dark spots denote greater values, shown in Fig. 4 [20, 21].

A gradient is a mathematical representation of an image's partials, which is a vector of its partials, denotes in Eq. (1) [22]

$$\nabla f = \begin{bmatrix} gx \\ gy \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \tag{1}$$

where $\frac{\partial f}{\partial x}$ the derivative at x direction, $\frac{\partial f}{\partial y}$ the derivative at y direction.

Central difference use to calculate convolution can be used to apply one-dimensional filter on an image to approximate the derivative image for finite differences, denotes in Eq. (2).

$$\frac{\partial f}{\partial y} = \begin{bmatrix} -1 \\ +1 \end{bmatrix} * A \tag{2}$$

where * represents the 1- convolution operation for dimensional. The 2 × 1 filter shifts the image a by half pixel. This may be avoided by employing the following (3 × 1) filter can be used, explain in Eq. (3):

$$\text{Filter}_{3\times 1} = \begin{bmatrix} -1 \\ 0 \\ +1 \end{bmatrix} \tag{3}$$

The formula below can be used to determine the gradient direction in Eq. (4a) [22]

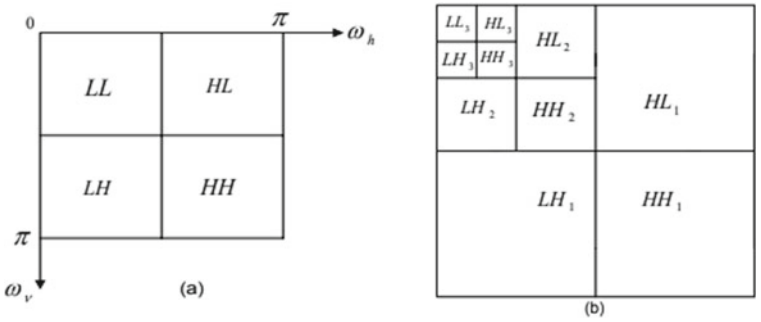


Fig. 5 **a** Location of the sub-bands in a two-dimensional spectrum (ω_h and ω_v denote the horizontal and vertical frequencies); **b** DWT with three levels of sub-band decomposition

$$\theta = \tan^{-1}\left(\frac{G_x}{G_y}\right) \tag{4a}$$

and the edge magnitude is given by Eq. (4b) [21]

$$G = \sqrt{g_y^2 + g_x^2} \tag{4b}$$

Discrete Wavelet Transform Technique (DWT)

Two types of wavelet transforms exist they are **(CWT)** continuous wavelet transform and (DWT) discrete wavelet transform [23]. This technique divides image signal into four distinct sub-bands as follows (LL, HL, LH, HH) indicating the type of the filters applied to the rows and columns of image shows in Fig. 5a, b, and Table 2 illustrates the example of transforms.

Discrete Cosine Transform Technique

This approach is frequently utilized in signal processing because it depicts a finite sequence of data points as a sum of cosine functions oscillating at various frequencies. There are eight conventional DCT versions, four of which are rather prevalent. Shows in Table 3 [34].

DCT provides high-pass bands with higher frequency accuracy and lower spatial resolution, while sub-bands provide the wavelet with high spatial accuracy and low frequency accuracy. It delivers a high compression ratio [1, 3] in DWT without sacrificing picture information, but it requires robust processing. DCT technology requires little processing, which implies that if any information is lost (containing artifacts) [35].

Table 2 Illustrates examples of transforms

Name of transform	Description
Haar wavelets	The input is a list of 2n elements; this approach is used to pair the values in the input, store the difference, and pass the total. This procedure is then repeated
Daubechies wavelets	In this technique, frequency relationships are used to generate progressively more accurate discrete samples for an implicit or wavelet function. And each time, the accuracy doubles from the previous accuracy [24–26]
On a dual-tree complex wavelet transform (DCWT)	The complex dual-tree wavelet transform (CWT) is a modification of the discrete wavelet transform (DWT), and its most important characteristic is that it is directional selective in two or more dimensions and is transformation constant. This is achieved with only the 2D repetition factor. Either the multidimensional double tree (M-D) CWT is not separable but depends on a computationally efficient and separable (FB) filter [27]
LeGall-Tabatabai (LGT)	Developed by Ali J. Tabatabai and Didier Le Gall in 1988 used in (JPEG XS or JPEG 2000) [28–30]
QMF with binomial distribution	Ali Naci Akansu created in 1990 [31]
(SPIHT) method for set partitioning in hierarchical trees	Developed in 1996 [32] by William A. Pearlman and Amir Said
Wavelet transform that is not decimated	(It is excluded from the down sampling)
Transform Newland	The orthonormal basis for wavelets is generated using correctly designed wavelets
Filters for top hats	In the frequency domain
Transforms of wavelet packets	In relation to (DWT) [33]
Wavelet transform that is complex	Is an alternative form

Contrast Technique

Contrast results from the interaction of the specimen with light, these include transmission, absorption, reflection, phase change, refraction, diffraction and fluorescence, and polarization (change in the wave vibration direction) [36]. Contrast has two types either point by point or subject by subject contrast, in point by point (deal with series of features of two subject) but in subject by subject contrast (discusses one subject and move to another); however, we can say contrast is the difference in color or luminance to distinguishable the object.

When the average brightness is high, tiny differences are unimportant, but when the average luminance is low (Weber–Fechner rule), little differences are significant. Then, the same small difference matters. Common definitions are given in Table 4.

Table 3 Four common standard of DCT [34]

DCT type	Equation description
DCT-I	$X_k = \frac{1}{2}\left(X_0 + (-1)^k X_{N-1}\right) + \sum_{n=1}^{N-2} X_n \cos\left[\frac{\pi}{N-1} nK\right] K = 0, \ldots, N-1$
DCT-II	$X_k = \sum_{n=0}^{N-1} X_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)K\right] K = 0, \ldots, N-1$
DCT-III	$X_k = \frac{1}{2} X_0 + \sum_{n=1}^{N-1} X_n \cos\left[\frac{\pi}{N} n\left(K + \frac{1}{2}\right)\right] K = 0, \ldots, N-1$
DCT-IV	$X_k = \sum_{n=0}^{N-1} X_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)\left(K + \frac{1}{2}\right)\right] K = 0, \ldots, N-1$
DCTV-VIII	DCTs I–IV process both limits in reference to a point of symmetry (even/odd), either at or near a data point. DCTs V–VIII signify (even/odd) words that are centered on a data point in one case and in the middle of two data points in the other

Table 4 Shows definition of (Weber, Michelson, RMS) contrast

Contrast name	Contrast equation	Description
Weber contrast	$\frac{I-I_b}{I_b}$	I and I_b illustrating the contrast between the brightness of the objects and the backdrop
Michelson contrast	$\frac{I_{\max}-I_{\min}}{I_{\max}+I_{\min}}$	$I_{\max}$ and $I_{\min}$ representing the highest and lowest luminance [37]. Michelson's anisotropy [38] (referred to as visibility) is a term that refers to patterns in which the bright and dark components are equal and create similar regions within an area (e.g., sine wave barriers)
RMS contrast	$\sqrt{\frac{1}{MN}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(I_{ij}-I)^2}$	RMS is the standard deviation of pixel intensity The root mean square (RMS) contrast is independent of the image's spatial contrast distribution or angular frequency content [39]

Co-occurrence Matrix Technique

The primary objective of this matrix is to count the number of times each ER appears in the same context as each EC to determine the context and the entities in which they co-occur [40].

Haralick defined the gray-level presence matrix as a square matrix that counts the number of times in an image that pairs of pixels with a specified value and offset appear [41].

For an image with distinct pixel values, the P × P co-occurrence matrix C is defined as follows across the $n \times m$ image I, with an offset $(\Delta x, \Delta y)$:

$$C_{\Delta x,\Delta y}(i, j) = \sum_{x=1}^{n}\sum_{y=1}^{m}\{1 \text{ if } I(x, y) = i \text{ and } I(x + \Delta x, y + \Delta y) = j\} \quad (5)$$

where i and j are the pixel values, x and y are the spatial positions in the image **I**, $I(x, y)$ indicates the pixel value at location (x, y) offsets, $(\Delta x, \Delta y)$ indicates the geographical relationship for which this matrix is being generated.

For each angle θ, the (contrast, correlation, energy, homogeneity) can be computed, Table 5 shows this formulas [42].

Fractal Technique

The fractal dimension is a measure of how "complex" a self-symmetric shape is. That is, it is a measure of the number of points in a given set. And all these three groups have the same number of points that is, each group is not countable, Table 6. Fractal dimension calculation [43–46].

Table 5 Formulas of co-occurrence technique [42]

Name of formula	Description
Contrast	$f^{\theta}_{\text{contrast}} = \sum_{i=1}^{L} \sum_{j=1}^{L} (i-j)^2 C_m$
Correlation	$f^{\theta}_{\text{correlation}} = \sum_{i=1}^{L} \sum_{j=1}^{L} C_m \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i)^2 (\sigma_j)^2}} \right]$ where $\mu_i = \sum_{i=1}^{L} \sum_{j=1}^{L} i C_m$, $\mu_j = \sum_{i=1}^{L} \sum_{j=1}^{L} j C_m$, $\sigma_i^2 = \sum_{j=1}^{L} \sum_{i=1}^{L} C_m \left(i - \mu_i^2\right)$, $\sigma_j^2 = \sum_{j=1}^{L} \sum_{i=1}^{L} C_m \left(j - \mu_j^2\right)$
Energy	$f^{\theta}_{\text{energy}} = \sum_{i=1}^{L} \sum_{j=1}^{L} C_M^2$
Homogeneity	$f^{\theta}_{\text{homogeneity}} = \sum_{i=1}^{L} \sum_{j=1}^{L} \frac{C_M}{1+\lvert i-j \rvert}$

Table 6 Fractal dimension calculation [47]

Equation	Description
$\text{dimension} = \frac{\log(\text{number of self similar pieccs})}{\log(\text{magnification factor})} = \frac{\log N^2}{\log N} = \frac{2 \log N}{\log N} = 2$	Dimension is the exponent of the number of self-similar fragments of the figure with magnification factor N that may be pulled apart As a result, logarithms are required. Take note that we have N^2 self-similar parts for the square, each with a magnification factor of N
$\text{dimension} = \frac{\log(\text{number of self similar pieccs})}{\log(\text{magnification factor})} = \frac{\log N^3}{\log N} = \frac{3 \log N}{\log N} = 3$	Similarly, a cube's dimension
$\text{fractal dimension} = \frac{\log(\text{number of self similar pieccs})}{\log(\text{magnification factor})}$	The definition of a self-similar object's fractal dimension
$\frac{\log(\text{number of self similar pieccs})}{\log(\text{magnification factor})} = \frac{\log 3}{\log 2} \approx 1.58$	To determine the size of S. A Sierpinski triangle is made up of three identical parts, each with a magnification factor of two
$\text{fractal dimension} = \frac{\log 9}{\log 4} = \frac{\log 3^2}{\log 2^2} = \frac{2 \log 3}{2 \log 2} = \frac{\log 3}{\log 2} \approx 1.58$	S dimension lies somewhere between 1 and 2 S also consists of nine self-contained pieces with a magnification factor of 4
$\text{fractal dimension} = \frac{\log 3^N}{\log 2^N} = \frac{N \log 3}{N \log 2} = \frac{\log 3}{\log 2}$	Similarly, S breaks into 3^N self-comparable bits with magnification factors 2^N

Principle Component Analysis (PCA) Technique

Principle component analysis (PCA) is a common method for feature extraction in data science. It reduces the number of features by creating numerical variants (new, smaller), and large part of information is captured that contained in the original features. The main steps in PCA as follows [48].

The first step is to equalize the ranges of continuous starting variables so that they contribute evenly to the analysis; if the ranges of initial variables vary significantly, those with larger ranges will dominate those with smaller ranges, resulting in skewed conclusions. This problem may be avoided by converting data to comparable scales.

Second step this stage aims to normalize the range of continuous prime variables in order to ensure that they all contribute equally to the analysis. If the ranges of the prime variables vary significantly, variables with greater ranges will dominate variables with lower ranges, resulting in skewed results. This issue may be avoided by converting data to comparable measures.

Third step eigenvalues and eigenvectors are linear algebra principles that are used to compute the primary components of data from the covariance matrix.

Fourth step here, the process of choosing between the important components is done, ignoring the less significant ones (low eigenvalues) and forming with the rest of the components an array of vectors that we call the feature vector.

Finally, the attribute vector is utilized to reroute data from the parent axes to the primary component-represented axes. This is accomplished by multiplying the original dataset's transposition by the transposition of feature vectors.

$$\text{Final Data Set} = \text{Feature Vector}^{\text{T}} * \text{Standardized Original Data Set}$$

3 Conclusion

In this study, the most important techniques such as (Gradient technique, discrete wavelet transform technique, discrete cosine transform technique, contrast technique, co-occurrence technique, and fractal and principal component analysis technique), they can be adopted in order to extract characteristics and analyzing the texture of the image for the purpose of their adoption in the classification of medical images were discussed. Applying these methods to RGB image and from these categories, we can propose a new mixed method for texture analysis and classification in precise ways.

References

1. Silverberg E, Boring CC, Squires TS (1990) Cancer statistics, 1990. CA Cancer J Clin 40:9–26
2. Siegel RL, Miller KD, Jemal A (2018) Cancer statistics, 2018. CA Cancer J Clin 68:7–30
3. Rogers HW et al (2015) Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the us population, 2012. JAMA Dermatol 151:1081–1086
4. Goodson AG, Grossman D (2009) Strategies for early melanoma detection: approaches to the patient with nevi. J Am Acad Dermatol 60:719–735
5. Argenziano G, Soyer HP, De Giorgi V (2000) Interactive atlas of dermoscopy. EDRA Medical Publishing & New Media, Milan, Italy
6. Pehamberger H, Steiner A, Wolff K (1987) In vivo epiluminescence microscopy of pigmented skin lesions; I pattern analysis of pigmented skin lesions. J Am Acad Dermatol 17:571–583
7. Argenziano G, Fabbrocini G, Carli P (1998) Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. Arch Dermatol 134:1563–1570
8. Menzies SW et al (2009) Dermoscopy: an atlas of surface microscopy of pigmented skin lesions. Sydney. McGraw-Hill, NSW, Australia
9. Menzies et al (1996) Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. Arch Dermatol 132:1178. PMID 8859028
10. Stolz W, Riemann A, Cognetta AB (1994) ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur J Dermatol 4:521–527
11. Henning JS et al (2007) The cash (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. J Am Acad Dermatol 56:45–52
12. Shaw HM, Rigel DS, Friedman RJ, Mccarthy WH, Kopf AW (2004) Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria. JAMA 92:2771–2776
13. Robinson JK, Turrisi R (2006) Skills training to learn discrimination of ABCDE criteria by those at risk of developing melanoma. Arch Dermatol 142:447–452
14. Korotkov K, Garcia R (2012) Computerized analysis of pigmented skin lesions: a review. Artifi Intell Med 56:69–90
15. Pathan S, Prabhu KG, Siddalingaswamy PC (2018) Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review. Biomed Sig Process Control 39:237–262
16. Barata C, Celebi ME, Marques JS (2019) A survey of feature extraction in dermoscopy image analysis of skin cancer. IEEE J Biomed Inform 23(3)
17. Irani M, Peleg S (1991) Improving resolution by image registration. Graph Models Image Process 53(3):231–239
18. Deepa SN, Sivanandam SN, Sumathi S (2006) Introduction to neural networks using MATLAB 6.0, vol 1. Tata McGraw-Hill
19. Finlayson GD, Qiu G, Qiu M (1999) Contrast maximizing and brightness preserving color to grayscale image conversion
20. Gradient Technique. http://www.google.com/search
21. Aggarwal A, Singh K, Singh K (2015) Use of gradient technique for extracting features from handwritten Gurmukhi characters and numerals. Procedia Comput Sci 46:1716–1723
22. Gonzalez R, Woods R (2008) Digital image processing, 3rd ed. Pearson Education, Inc, Upper Saddle River, New Jersey. ISBN 978-0-13-168728-8
23. Discrete Wavelet Transform Technique. https://www.sciencedirect.com› science › article › pii
24. Akansu AN, Haddad RA, Caglar H (1990) Perfect reconstruction binomial QMF-wavelet transform. In: Proceedings of SPIE visual communications and image processing, Lausanne, vol 1360, September 1990, pp 609–618
25. Akansu AN, Haddad RA (1992) Multiresolution signal decomposition: transforms, subbands, and wavelets. Academic Press, Boston, MA. ISBN 978-0-12-047141-6
26. Akansu AN (1993) Filter banks and wavelets in signal processing: a critical review. In: Proceedings of SPIE video communications and PACS for medical applications (Invited Paper), vol 1977, Berlin, October 1993, pp 330–341

27. Selesnick IW, Baraniuk RG, Kingsbury NC (2005) The dual-tree complex wavelet transform
28. Sullivan G (2019) General characteristics and design considerations for temporal subband video coding, 8–12 December 2003. ITU-T. Video Coding Experts Group. Retrieved 13 September 2019
29. Bovik AC (2009) The essential guide to video processing. Academic Press, p 355. ISBN 9780080922508
30. Gall DL, Tabatabai AJ (1988) Sub-band coding of digital images using symmetric short kernel filters and arithmetic coding techniques. In: ICASSP-88, International conference on acoustics, speech, and signal processing, vol 2, pp 761–764. https://doi.org/10.1109/ICASSP.1988.196696, S2CID 109186495
31. Daubechies wavelets. In: Proceedings of 1st NJIT symposium on wavelets, April 1990
32. Said A, Pearlman WA (1996) A new, fast, and efficient image codec based on set partitioning in hierarchical trees. IEEE Trans Circ Syst Video Technol 6(3):243–250. https://doi.org/10.1109/76.499834. ISSN 1051-8215. Retrieved 18 October 2019
33. FastWaveletTransform. https://en.wikipedia.org/w/index.php?title=Discrete_wavelet_transform&action=edit&section=6
34. DiscreteCosinetransform https://en.wikipedia.org/wiki/Discrete_cosine_transform
35. Katharotiya A, Patel S, Goyani M (2011) Comparative analysis between DCT and DWT techniques of image compression. J Inform Eng Appl 1(2). ISSN 2224-5758 (print). ISSN 2224-896X (online)
36. Contrast Technique. http://www.microscopist.co.uk/essential-techniques/contrast-techniques
37. Arend L (2018) Luminance contrast. Color usage.arc.nasa.gov. Retrieved 5 April 2018
38. Michelson A (1927) Studies in optics. U. of Chicago Press
39. Peli E (Oct 1990) Contrast in Complex Images (PDF). J Opt Soc America A. 7(10):2032–2040. https://doi.org/10.1364/JOSAA.7.002032. Archived from the original (PDF) on 2016-05-21. Retrieved 2009-02-16
40. Texture analysis using the gray-level co-occurrence matrix (GLCM)—MATLAB and simulink—MathWorks United Kingdom. uk.mathworks.com. Retrieved 2020-06-26
41. Nanni L, Brahnam S, Ghidoni S, Menegatti E, Barrier T (2013-12-26) Different approaches for extracting information from the co-occurrence matrix. PLOS One 8(12):e83554. Bibcode: 2013PLoSO...883554N, https://doi.org/10.1371/journal.pone. 0083554. ISSN 1932-6203. PMC 3873395.PMID 24386228
42. Bhargava D, Bansal A (2020) Comparative analysis of classification techniques for brain magnetic resonance imaging images. In: Advances in computational techniques for biomedical image analysis
43. Mandelbrot BB (1983) The fractal geometry of nature. Macmillan. ISBN 978-0-7167-1186-5
44. Wiley. xxv, Falconer, K (2003) Fractal geometry: mathematical foundations and applications. ISBN 978-0-470-84862-3
45. Briggs J (1992) Fractals: the patterns of chaos. Thames and Hudson, London, p 148. ISBN 978-0-500-27693-8
46. Vicsek T (1992) Fractal growth phenomena, vol 31. World Scientific, Singapore/New Jersey, pp 139–146. ISBN 978-981-02-0668-0
47. Fractal technique. http://www.sciencedirect.com>science>artical>pi
48. Principal Component Analysis Technique. https://iq.opengenus.org› algorithm-principal-component analysis

Identification of Key Criteria of Selecting the Delivery System and Type of Contract in Construction Projects

Sajjad Ali Mahmood Alkaabi and Ahmed Mohammed Raoof Mahjoob

Abstract Construction projects are becoming more complicated, and there are a growing number of elements that contribute to project failure. As a result, a more systematic approach to selecting an appropriate project delivery system and kind of contract is required. The selection of a construction project delivery system and type of contract greatly influences the construction project outcome and considers one of the most important factors of the construction project success. So, the present study aims to identify the key criteria of selecting the project delivery system and also identify the key criteria of selecting the type of contract of the construction project. This study used the Delphi survey and principal components analysis to achieve this aim. The results of this study showed that there are Thirteen key criteria for selecting the construction project delivery system, where the criterion of "Clarity of contractual treatments" was of the highest importance with a value of Mean is (4.334). In addition, this study identified Eleven key criteria for selecting the type of contract in the construction project, where the criterion of "Bear the risk of unexpected cost increases" was of the highest importance with a value of Mean is (4.5). While the results of applying principal components analysis on the Delphi survey outcomes showed that these criteria are most valuable on the decision-making process and there is no possibility to reduce its number.

Keywords Delivery system · Selection criteria · Type of contract · Delphi survey · Principal components analysis

1 Introduction

The usage of the correct project delivery system (PDS) is one of the most important components in attaining the specified goals of the construction project [1]. The right

S. A. M. Alkaabi (✉) · A. M. R. Mahjoob
University of Baghdad, Baghdad, Iraq
e-mail: s.mahmood1901p@coeng.uobaghdad.edu.iq

A. M. R. Mahjoob
e-mail: Ahmed.mahjoob@coeng.uoBaghdad.edu.Iq

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_14

PDS can help prevent problems and achieve project-specific goals like quick project completion, low cost, practical risk allocation between the parties, and giving the owner control over the details of the design solution and the amount of in-house work, among other things [2]. In addition, choosing the right PDS might help boost the project's results [3]. A project delivery system is a kind of contracting in which the owner's risk and responsibility for design and construction are transferred to another party. Normally, a design entity is in charge of the design, while a contractor is in charge of the performance of construction [4]. In the public sector, this has virtually always meant employing the design-bid-build system, which entails the separation of design and construction services as well as sequential design and construction performance [5]. The choice of a PDS is based on a number of criteria. The market environment may have an impact on the selection of the PDS in addition to project properties, the owner's objectives, and available know-how and resources [2, 4]. Inappropriate PDSs are chosen based on little project knowledge, biased prior experience, and a poorly defined, if any, list of factors to evaluate [6, 7]. There are three areas consider very important issues for the owner in the construction projects and need to owner decision [8]:

- Project Delivery System such as (Design-Build, Design-Bid-Build, and Construction Management). (This study area).
- Procurement Method (Best Value, Low Bid, Negotiated, Qualifications-Based, and Sole Source).
- Contract Type (Unit Price, Target Price, Lump Sum, and Cost-Plus Fee). (This study area).

The design phase, tender phase, and construction phase are all part of the procurement process for any construction project. However, in terms of cost, the most essential contractual relation is that between the Owner and the Contractor. As a result, the Contract Type selected, which specifies the manner of payment, is critical in ensuring that construction costs remain within the approved funding constraints [9]. After identifying the PDS, the owner faces other challenge that is choosing the right construction contract type. There are a few studies that focus on identifying the selection criteria of PDS than the selection method of PDS. This study aims to identify the key criteria of selection the project delivery system by using Delphi survey with principal components analysis (PCA) in Iraqi construction projects. Also, it is aims to identify the selection criteria of construction contract type by using Delphi survey with PCA in Iraq construction projects.

2 Research Method

Most authors focused on the methods of selection and their accuracy but did not give identifying the selection of criteria adequate attention in their studies, especially when selecting the construction contract type and the project delivery system, so the purpose of the present study is:

1. To identify the key selection criteria of the project delivery system for the Iraq construction projects by using the Delphi survey and principal component analysis (PCA).
2. To identify the key selection criteria for selecting the construction contract type for the Iraq construction projects by using the Delphi survey and principal component analysis (PCA).

The approach of this study is a qualitative approach. The literature review and Delphi survey were used to collect the needed data. For statistical analysis for the gathered information from the Experts in Delphi survey, this study used The XLSTAT software (Version 2015). The present study was conducted in Baghdad, Iraq.

3 Results and Discussions

3.1 Conducting the Delphi Rounds to Identify the Selection Criteria

After reviewing the related literature, the research visited the related Iraq government directorates, like (Al-Mansour General Engineering Company, Housing Directorate, Buildings Directorate, etc.) were visited to gather more information about the selection criteria of PDS and type of contract for the construction projects. Then, the Delphi survey with group of experts' engineers were conducted to identify the key selection criteria of both the PDS and type of contract for the construction projects. The process of Delphi survey depends on developing the questionnaire form for each round. The research, like [10] sent Twenty-one invitations to ask experts to participate in the Delphi rounds, all the selected sample of experts having experience more than fifteen years in supervision, management and construction of the Iraq construction projects. The invitation explained the purpose of the study and a number of rounds, and the method of the distribution of the questionnaire form. Twelve experts responded and agreed to participate and the 75% they from public work sector, the academic degrees, and field of specialization for the experts are given in Figs. 1 and 2, respectively.

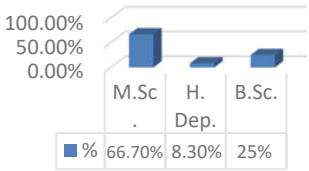


Fig. 1 Experts academics degree

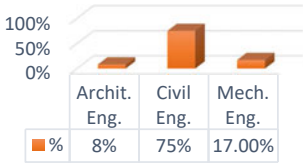


Fig. 2 Field of specialization for experts

Table 1 Weight value (WV) of descriptive frequencies [11]

Descriptive frequency	Class interval	Weight value (WV)
Very low	1 ≤ Class interval ≤ 1.8	1
Low	1.8 < Class interval ≤ 2.6	2
Medium	2.6 < Class interval ≤ 3.4	3
High	3.4 < Class interval ≤ 4.2	4
Very high	4.2 < Class interval ≤ 5	5

Delphi First Round

The questionnaire form of first round of the Delphi divided to two parts (part one includes the selection criteria of PDS, while the part two includes the selection criteria of contract type) which formulated according to literature review and field visits, like [11] the research used five-Likert scale with the weight value that shown in Table 1. Then questionnaire form distributed to the twelve experts and they were asked to identify the importance degree for each criterion in the both parts with the possibility to add other criteria by the participants if they are considered it necessary. In this round a face-to-face interview with each participate were conducted to ensure the correct answers.

The answers and suggestions by the experts in this round were analyzed carefully. The research used the arithmetic mean, standard deviation (SD), and Cronbach alpha (α) for the statistical analysis by using XLSTAT software (Version 2015). The analysis results showed that the average value of Cronbach alpha (α) for all criteria was more than (0.7), the criteria that got arithmetic mean less than number (3.4), were identified and will be removed it in second round of the Delphi survey, also according to participants views there are a number of criteria were merged, removed, and added so new lists of criteria were formed for both parts. So, the list of selection criteria of PDS became includes fifteen criterion, while the list of selection criteria of the contract type became includes twelve criterion. At the end of the first round these lists of criteria will be used in the second round of Delphi survey.

Delphi Second Round

In this round, the experts were asked to indicate the relevant importance for the selection criteria of PDS and contract type that resulted from the Delphi first round. The statistical analysis for the experts' answers were conducted in this round, the analysis results showed that there are thirteen selection criteria of PDS resulted from this round and will use in the Delphi third round, and only two criteria got arithmetic mean less than (3.4) will be remove from the Delphi third round. Also, the analysis results showed that the average value of (α) for these resulted criteria was (0.904), this indicates that experts' answers have a high reliability. While there eleven selection criteria of construction contract type resulted from this round and will use in the Delphi third round, only one criterion got arithmetic mean less than (3.4) will remove from the Delphi third round. The average value of (α) for the selection criteria of contract type that resulted was (0.897), this indicates that experts' answers have a high reliability also.

Delphi Third Round

In this round the same experts were asked to re-evaluate the relative importance of the selection criteria of PDS and the selection criteria of construction contract type that passed from the second Delphi round, to find the key selection criteria that affect the decision-making process in Iraq construction projects. The results of the statistical analysis in this round showed that there are thirteen selection criteria of PDS got an arithmetic mean more than (3.4) and the calculated (α) for these selection criteria was (0.908). While there are eleven selection criteria of construction contract type got an arithmetic mean more than (3.4) and the calculated (α) for these selection criteria was (0.939). So, these results are similar to the results of the Delphi second round, this indicates that there is a consistency in experts' answers in this round and there is no need to go to the fourth round of Delphi survey, Table 2 illustrates the final selection criteria of PDS from Delphi survey, while Table 3 illustrates the final selection criteria of construction contract type from Delphi survey.

Principal Component Analysis (PCA)

The principal component analysis (PCA) were conducted by using XLSTAT software (Version 2015) on the experts' answers in the third Delphi round to find the possibility of reducing the selection criteria for both PDS and construction contract type that identified in the Delphi third round, Dimensionality reduction is one of the key applications of principal component analysis (PCA), which has been widely employed by researchers in a variety of fields [12].

The PCA extracts the most valuable features with a factor pattern coefficient of more than (0.3) [13]. The Kaiser–Meyer–Olkin (KMO) test must be completed before the PCA may be performed. KMO can tell whether the data is accepted to do the PCA. When the value of KMO is less than (0.5) this refers to unaccepted data to run the PCA [14]. The research used XLSTAT software (Version 2015) to calculate the KMO value and PCA, the results showed that the value of KMO for the selection criteria of PDS was (0.528), after applying the PCA, the results showed

Table 2 The final selection criteria of PDS

No	Selection criteria	Mean	SD
1	Clarity of contractual treatments	4.3333	0.65134
2	Type and complexity of the construction project	4.3333	0.88763
3	Available time	4.3333	0.98473
4	Contractor capacity and availability of competitors	4.2500	0.62158
5	Availability of technology and potential for innovation	4.2500	0.86603
6	Required quality	4.1667	0.83485
7	Owner's ability to take risks (risk distribution)	4.1667	0.93744
8	Life cycle cost	4.0833	0.79296
9	Owner's experience	4.0833	0.99620
10	Management and coordination strength of the owner	4.0000	0.60302
11	Flexibility in design idea	4.0000	0.60302
12	Resources available to the owner	4.0000	0.85280
13	The possibility of fast track	3.9167	0.79296

Table 3 The final selection criteria of construction contract type

No	Selection criteria	Mean	SD
1	Bear the risk of unexpected cost increases	4.5000	0.79772
2	Commitment to the set schedule	4.5000	0.90453
3	Employer's staff experience	4.5000	0.90453
4	Expected project costs	4.4167	0.66856
5	Work quality guarantee	4.4167	0.90034
6	The price competition	4.3333	0.77850
7	Project type and complexity	4.3333	1.15470
8	Availability of resources	4.2500	0.86603
9	Contractual difficulties and Claims	4.1667	0.57735
10	Uncertainty about the scope of work required	4.0000	0.73855
11	The possibility of fast track	3.9167	0.66856

that there are two components that have eigenvalue more than (1) will depend on these components to extract the factor pattern because of these components gives a representation of the optimal solution with a rate more than (80%). The values of KMO and the values of factor pattern are illustrated in Table 4. So, all the identified criteria are most valuable and there is no possibility to reduce because the factor pattern for all the selection criteria was more than (0.3). While the value of KMO for the selection criteria of construction contract type was (0.481) this result refers to unacceptable data to run the PCA because the value of KMO is less than (0.5). The final selection criteria identified from the Delphi survey stayed without any reduction after applying the PCA.

Table 4 The values of KMO and the factor pattern of PCA for the selection criteria of PDS

Selection criteria of the PDS	KMO	Factor pattern	
		F1	F2
1. Clarity of contractual treatments	0.506	**0.910**	−0.089
2. Type and complexity of the construction project	0.480	**0.868**	0.214
3. Available time	0.984	**0.777**	0.184
4. Contractor capacity and availability of competitors	0.397	**0.735**	0.137
5. Availability of technology and potential for innovation	0.507	**0.818**	0.057
6. Required quality	0.677	**0.933**	−0.226
7. Owner's ability to take risks (risk distribution)	0.543	**0.940**	0.217
8. Life cycle cost	0.588	**0.926**	−0.275
9. Owner's experience	0.765	0.553	**0.780**
10. Management and coordination strength of the owner	0.394	**0.766**	−0.331
11. Flexibility in design idea	0.536	**0.915**	0.290
12. Resources available to the owner	0.455	**0.871**	−0.337
13. The possibility of fast track	0.487	**0.852**	−0.333
The final value of KMO	0.528		

Values in bold correspond for each variable to the factor for which the squared cosine is the largest

4 Conclusion

After reviewing the relevant literature on the project delivery systems and their selection criteria, and the construction contract types and their selection criteria, there are very few studies that outlined the selection criteria and the methods of identifying them were founded, in spite of the fact that their high affect success the construction project. So, the present study used three rounds from the Delphi survey with applying the principal components analysis (PCA) to identify the key selection criteria of the project delivery system (PDS) and also to identify the key selection criteria of construction contract type in order to facilitate the decision-making process for the owner or his representative. The results of this study showed that there are thirteen key selection criteria of PDS in Iraq construction projects, where the criterion of "Clarity of contractual treatments" is the most important in the selection process. While there are eleven key selection criteria of construction contract type for the Iraq construction projects, where the criterion of "Bear the risk of unexpected cost increases" is the most important in the decision-making process. Also, the select panel of experts in the Delphi survey was suitable and reflected the good views as shown in PCA results. Finally, these selection criteria will help the decision-makers to ensure effective decisions on selecting the suitable PDS and the construction contract type in Iraq construction projects.

Acknowledgements This study was created in collaboration with the University of Baghdad's Department of Civil Engineering and the Ministry of Housing and Construction/Directorate of Buildings.

Conflict of Interest
The corresponding author declares that there is no conflict of interest on behalf of all authors.

References

1. Chen YO, Liu JY, Li B, Lin B (2011) Project delivery system selection of construction projects in China. Exp Syst Appl 38:5456–5462
2. Lahdenperä P (2015) Project delivery systems in Finnish new building construction—a review of the last quarter century. In: 8th Nordic conference on construction economics and organization. Elsevier, Finland, pp 162–169
3. Hong HK, Kim JS, Kim T, Leem BH (2008) The effect of knowledge on system integration project performance. Ind Manag Data Syst 108(3):385–404
4. Mahdi IM, Alreshaid K (2005) Decision support system for selecting the proper project delivery method using analytical hierarchy process (AHP). Int J Project Manage 23:564–572
5. Trauner CS (2007) Construction project delivery systems and procurement practices: considerations alternatives advantages disadvantages. Trauner Consulting Services, Inc
6. Rwelamila PD, Edries R (2007) Project procurement competence and knowledge base of civil engineering consultants: an empirical study. J Manag Eng 23:182–192
7. Qiang M, Wen Q, Jiang H, Yuan S Factors governing construction project delivery selection: a content analysis. Int J Proj Manag. https://doi.org/10.1016/j.ijproman.2015.07.001
8. Design-build done right primer. Design-Build Institute of America Publication (2015)
9. Antoniou F, Aretoulis GN, Konstantinidis D, Kalfakakou GP (2012) Selection criteria used for the choice of contract type for major highway construction projects. In: Europe 2012 conference on transport research arena (TRA). Elsevier, Athens, pp 3508–3517
10. Al-Agele HKB, Al-Kaabi SAM (2016) Identification of key factors affecting waste management in life cycle of the construction project by using Delphi technique. J Eng 22(7):19–34. Available at https://joe.uobaghdad.edu.iq/index.php/main/article/view/187
11. Mohammed SR, Jasim AJ (2018) Examining the values and principles of agile construction management in Iraqi construction projects. J Eng 24(7):114–133. https://doi.org/10.31026/j.eng.2018.07.08
12. Preetha R, Vinila JS (2020) Early diagnose breast cancer with PCA-LDA based FER and neuro-fuzzy classification system. J Ambient Intell Humanized Comput. Available at https://doi.org/10.1007/s12652-020-02395-z
13. Lever J, Krzywinski M, Altman N (2017) Points of significance: principal component analysis. Nat Methods 14(7):641–642
14. Mahmoudi A, Deng X, Javed SA, Yuan J (2020) Large-scale multiple criteria decision-making with missing values: project selection through TOPSIS.OPA. J Ambient Intell Humanized Comput. https://doi.org/10.1007/s12652-020-02649-w

Design and Implementation of a Low-Cost Weather Stations Meter

Israa S. Al-Furati, Fatemah K. Al-Assfor, and Atheel K. Abdul Zahra

Abstract The weather monitoring system is used to provide pharmacists, farmers, event planners, and others with a precise statistic to guide them in taking an appropriate action. Today, with the last increased in smart technologies, the system is developed in too many sensing techniques to capture the real-time climate data for a wide area. In this paper, the proposed system is used to calculate the weather (hot, cold, wind speed, humidity, temperature, and dry) in an indoor and outdoor environment in order to monitor and control the weather conditions. This system is simple and easy to implement which consists of two parts: The first part is used only in the outdoor environment, and the second one is used in an indoor and outdoor environment. A very simple hardware components are used to build the whole system such as the DHT-11 sensor, LCD (16×4) screen, Arduino microcontroller, and an encoder to calculate the wind speed. The system can be extended to control different tasks based on weather conditions. The results are summarized in a table and sent to the control unit system.

Keywords Arduino microcontroller · LCD display · Wind speed · Weather station system · Temperature and humidity sensor

1 Introduction

In the last decade, the weather conditions (humidity and temperature) are very important issue for human because it make them updated with any frequent weather condition [3]. Usually, the weather station in our daily life is built in a very large size (bulk) and consumes a more power conception (see Fig. 1) which offers a limited reliability [4, 5]. In a spatial case, the need is going to design a small system with

I. S. Al-Furati (✉)
Department of Electrical Engineering, University of Basrah, Basrah, Iraq
e-mail: israa.sabri@uobasrah.edu.iq

F. K. Al-Assfor · A. K. Abdul Zahra
Department of Computer Engineering, University of Basrah, Basrah, Iraq

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_15

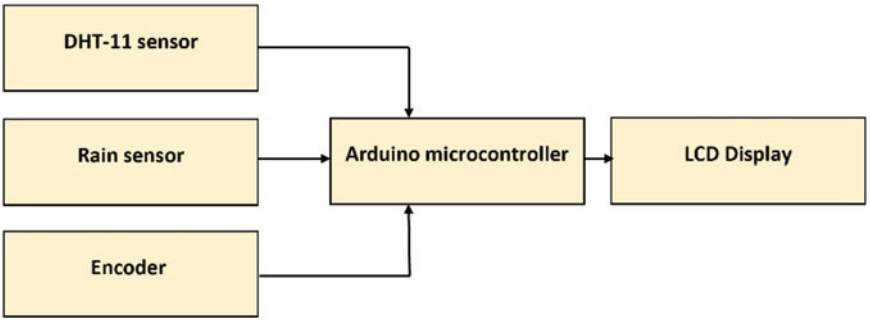


Fig. 1 Weather station module for outdoor environment

low-power conceptions and portable in order to measure the various conditions environments. Furthermore, it is efficient in remote places. To cover this requirement, we suggest this system to calculate the weather (hot, cold, wind speed, humidity, temperature, and dry) in an indoor and outdoor environment in order to monitor and control the weather conditions. This system is simple and easy to implement which consists of two parts: The first part is used only in the outdoor environment, and the second one is used in an indoor and outdoor environment. A very simple hardware components are used to build the whole system such as the DHT-11 sensor, LCD (16×4) screen, Arduino microcontroller, and an encoder to calculate the wind speed. The android applications are used to falcate the task by enable the user to reaching all the measurements using only on tap [6]. Also, the results are monitored in more details through the humidity and temperature sensors controlled by the user. In many systems, a Bluetooth is used to transmit the data which are obtained using the sensors in the environment through the android application in order to notify the user about the weather condition through the mobile phone application [7, 8]. The aim of this system is to propose a small portable system with low cost, simple, easy to implement capability that notify us about the status of the outdoor temperature, daily amount of rainfall humidity, wind speed barometric pressure. This system gives us the information about the weather conditions through the mobile phone using simply the android application. Figure 1 shows the flowchart that describe the suggested system. The rest of the paper is organized as follows: Sect. 2 describes the outdoor meter system; Sect. 3 describes the indoor and outdoor meter system; Sect. 4 explains the result obtained from this system and its discussion, and Sect. 5 shows the conclusion.

2 Outdoor Weather Meter System

The proposed system (see Fig. 2) is used to calculate the weather (hot, cold, wind speed, humidity, temperature, and dry) in an outdoor environment in order to monitor and control the weather conditions. This system is easy to implement and a very

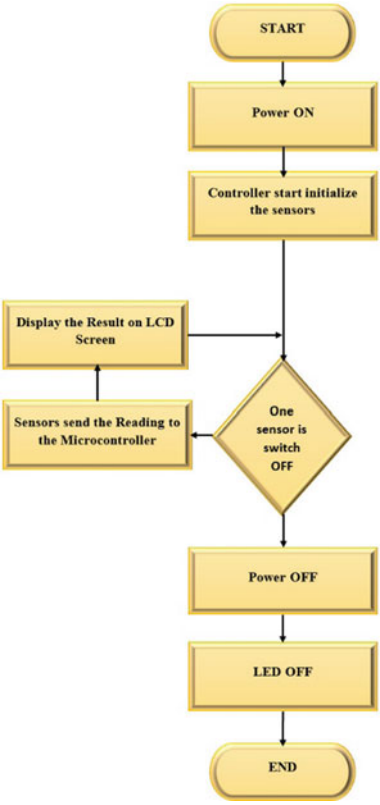


Fig. 2 Flowchart of the working principles

simple hardware components such as the DHT-11 sensor [9, 10] which is used to measure the temperature and humidity, rain sensor that is used to sense the rain in the weather and an encoder connected with a fan to calculate the wind speed [11, 12]. All these components are connected through an Arduino microcontroller to display the result on the LCD (16×4) screen as shown in Fig. 3. When the power is switched ON, the sensors initialize reading the weather status and send the results through the Arduino Uno microcontroller to the A-TO-C IC module that is connected to the LCD screen to display the weather conditions. The suggested box is a small and simple with easy to implement property. The initialization is started by push ON the switch, and the LED is lighted, and when the operation is ended, the switch is push OFF and the LED is OFF.

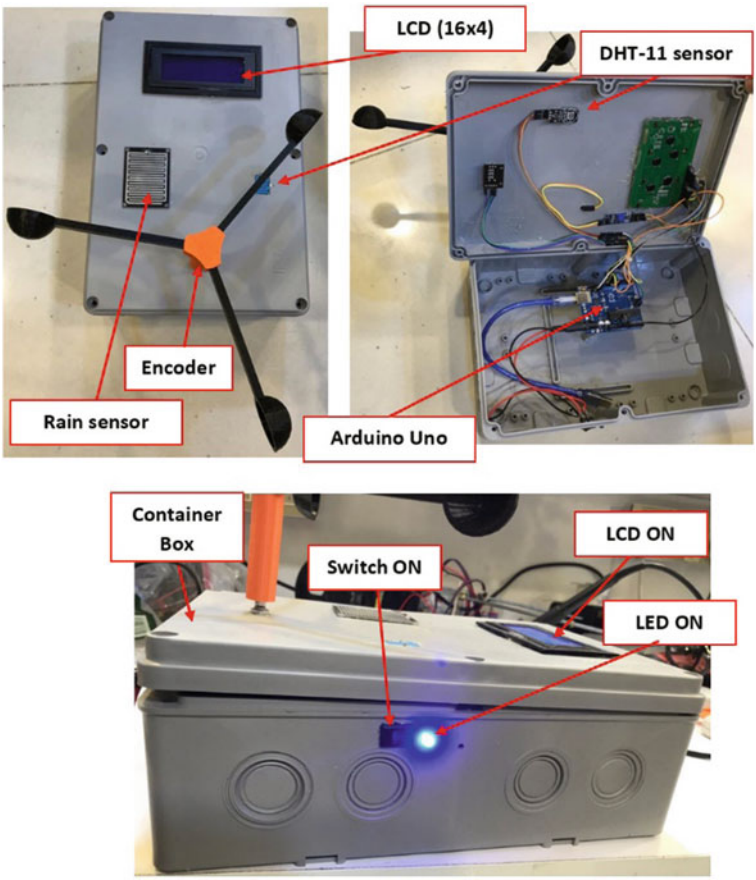


Fig. 3 Outdoor meter system

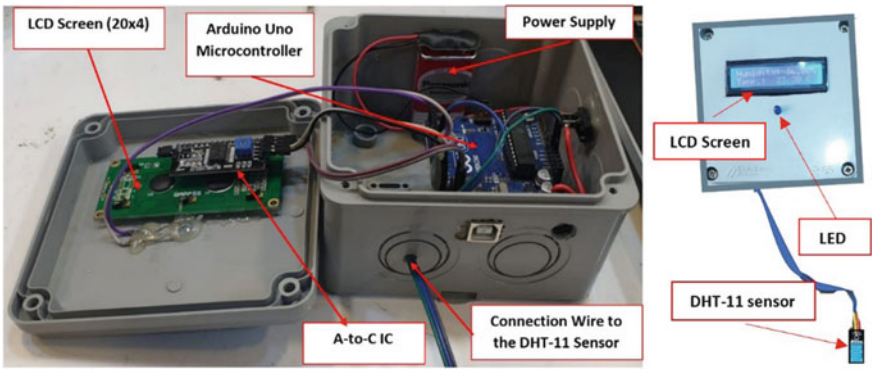


Fig. 4 Indoor and outdoor meter system

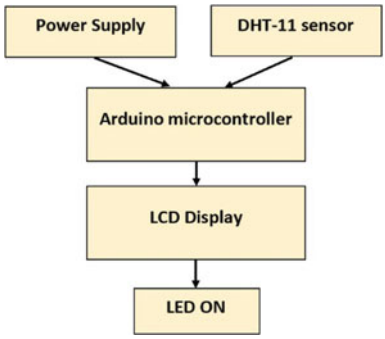


Fig. 5 Weather station module for indoor and outdoor environment

3 Indoor and Outdoor Weather Meter System

This system (see Fig. 4) is used to calculate the weather (humidity and temperature) in an indoor and outdoor environment in order to monitor and control the weather conditions. This system is built using the DHT-11 sensor to measure the temperature and humidity which is connected through an Arduino microcontroller to display the result on the LCD (16×4) screen as shown in Fig. 5.

4 Result and Discussion

In this paper, two-meter box is suggested to calculate the whole weather conditions in an indoor and outdoor environment. In the first system (see Fig. 2), when the push button switch is ON, the system is used to calculate the weather conditions. Table 1 is a sample of the weather conditions in only one day at different times with different climates in the city of Basra. The results are summerized in Table 1. It is found that in the Monday, the range of temperature is about (31–42); the range of the wind speed is about (14–25) in the morning, (22–25) in the afternoon, (9–14) in the evening, and (18–25) in the night while the humidity about (9–20) in the morning, (7–9) in the afternoon, (9–14) in the evening, and (15–19) in the night. Rain case is occurred only when the humidity is high and the prevailing climate is dry. For the second system (see Table 2) in the outdoor, there are two conditions to be calculated which are the result obtained in the sun and the shadows. It is found that the range of temperature is similar to that obtained by the first system and its range about (30–43), and the humidity is about (9–22)%. Furthermore, it is found that the temperature under the sun is more than the temperature under the shadow, while the humidity under the shadow is more than the humidity under the sun. By studying the result in Tables 1 and 2, it is found that the temperature is decreased as the time is late (in the night), and it is in a high temperature when the time is in the mid (afternoon) as shown in Fig. 6, while the humidity is simulated in Fig. 7. It is clear that the relation is inverse between temperture and humidity, whereas, the time is going late the humidity is

Table 1 Weather condition in several area

Locations	Days	Time	Weather Conditions			
			Temperature	Humidity (%)	Wind Speed (KM/II)	Rain
Zubair	Monday	Morning	32	9	14	Dry
		Afternoon	42	7	22	Dry
		Evening	37	10	9	Dry
		Night	35	19	19	Dry
Najebea	Monday	Morning	31	12	20	Dry
		Afternoon	39	6	24	Dry
		Evening	36	12	11	Dry
		Night	34	15	25	Dry
Qurna	Monday	Morning	32	20	25	Rain
		Afternoon	37	8	25	Dry
		Evening	37	9	14	Dry
		Night	33	14	21	Dry
Basrah city	Monday	Morning	33	17	21	Dry
		Afternoon	40	9	23	Dry
		Evening	38	14	12	Dry
		Night	35	15	22	Rain

increase since the temperature is decrease and so on. On the other hand, the wind speed is calculated by the encoder connected through the fan and result in the wind speed. The variation is between (0–40) that is mean: The air speed varies between 0 and the 40 as shown in Fig. 8. At last, Fig. 9 shows the overall result of all the sensors that is connected in this system, and also, the maximum wind speed is also calculated and summarized in Fig. 9.

5 Conclusion

The suggested system is used to provide the meteorological with a precise statistic to guide them in taking an appropriate action and control the weather conditions. This system is simple and easy to implement with a very simple hardware components which consists of two parts: The first part is used only in the outdoor environment, and the second one is used in an indoor and outdoor environment. The results are summarized in a table and sent to the control unit system. Experimental results show high sensitivity variation to the temperature and humidity which can be easily calibrated and controlled by C++ code on the Arduino Uno IDE. The measured temperature and humidity are presented numerically on a small LCD screen connected to the Arduino

Table 2 Weather condition in several

Locations	Days	Time	Weather Conditions			
			Temperature (Under shadows)	Temperature (Under sun)	Humidity (%) Under shadows	Humidity (%) Under sun
Outdoor	Mon.	Morning	31	33	14	10
		Afternoon	40	43	22	12
		Evening	37	37	9	22
		Night	33	32	19	9
Outdoor	Tue.	Morning	30	32	16	8
		Afternoon	39	40	20	12
		Evening	37	38	12	9
		Night	32	34	10	16
Outdoor	Wed.	Morning	34	36	12	10
		Afternoon	39	40	21	8
		Evening	37	38	10	12
		Night	34	37	17	9
Outdoor	Thu.	Morning	32	35	17	12
		Afternoon	39	40	21	8
		Evening	37	38	10	10
		Night	36	35	14	8
Outdoor	Fri.	Morning	30	34	16	13
		Afternoon	33	38	21	10
		Evening	30	37	13	10
		Night	30	32	18	15
Outdoor	Sat.	Morning	37	39	16	10
		Afternoon	33	38	21	12
		Evening	35	36	13	9
		Night	32	30	18	10
Outdoor	Sun.	Morning	30	32	16	8
		Afternoon	39	40	20	12
		Evening	37	38	12	9
		Night	32	34	10	16

DO (Digital Output) ports. Thus, the sampling time and measurement accuracy are adjusted directly in the in C ++ program.

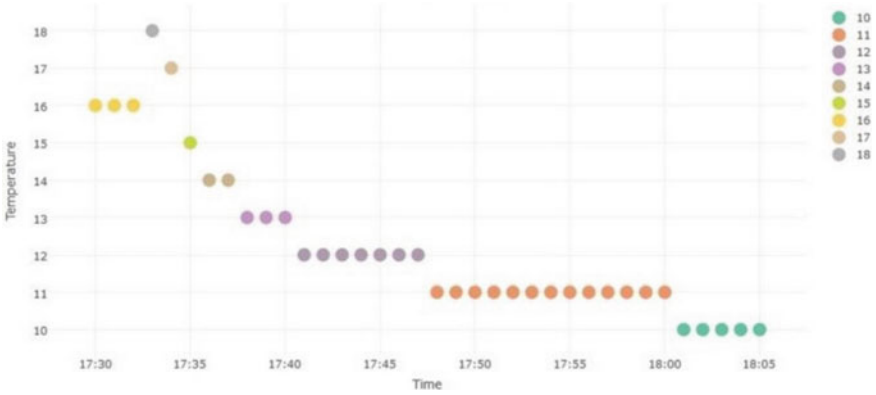


Fig. 6 Temperature result by the DHT-11 sensor

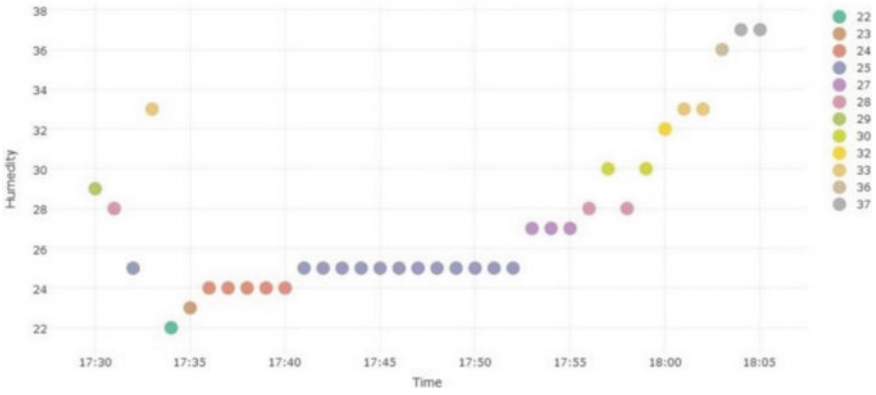


Fig. 7 Humidity result by the DHT-11 sensor

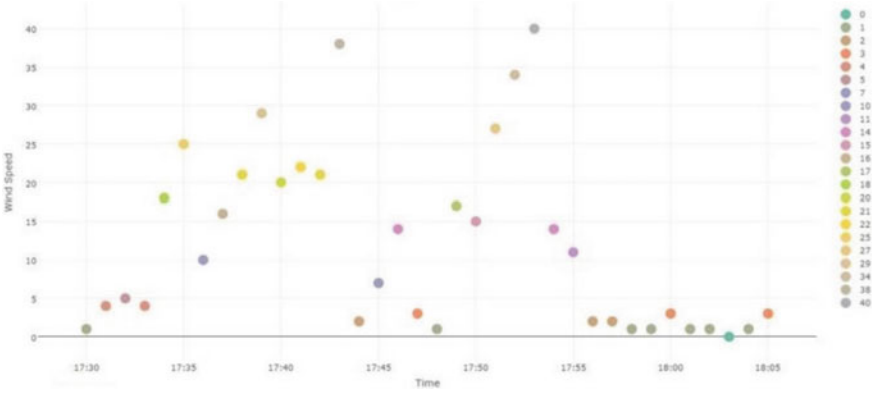


Fig. 8 Wind speed result by the encoder

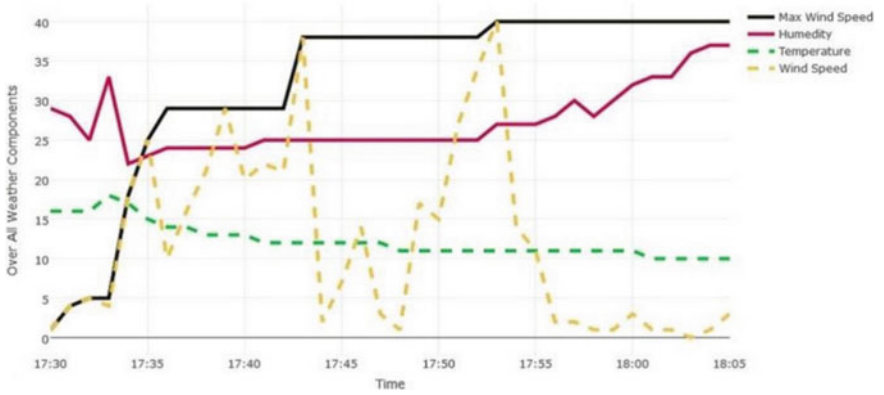


Fig. 9 Weather station components

References

1. Al-Furati IS, Rashid AT (2021) Design and implementation of indoor robot positioning system using LED array and LDR sensor. J Eng Sci Technol 16(2):1688–1710
2. Gangopadhyay S, Mondal MK (2016) A wireless framework for environmental monitoring and instant response alert. In: Microelectronics, computing and communications (MicroCom), Durgapur, India
3. WeatherFlow (2017) Smart home weather stations. http://weatherflow.com/smart-home-weather-stations/, Son Eriúim Tarihi: 30 Haziran 2017
4. Malhotra A, Som S, Khatri K (2019) IoT based predictive model for cloud seeding. In: Amity international conference on artificial intelligence (AICAI), pp 669–773
5. Karra A, Kondi B, Jayaraman R (2020) Implementation of wireless communication to transfer temperature and humidity monitoring data using Arduino Uno. In: 2020 international conference on communication and signal processing (ICCSP), pp 1101–1105
6. Kaneriya S, Tanwar S, Buddhadev S (2018) A range-based approach for long-term forecast of weather using probabilistic Markov model. In: 2018 IEEE international conference on communications workshops (ICC Workshops), pp 1–6
7. Al-Furati IS, Rashid AT (2020) Shortest distance orientation algorithm for robot path planning using low-cost IR sensor system. In: 2020 international conference on electrical, communication, and computer engineering (ICECCE). IEEE, pp 1–6
8. Benghanem M (2010) RETRACTED: a low-cost wireless data acquisition system for weather station monitoring, pp 862–872
9. Munandar A, Fakhrurroja H, Rizqyawan MI, Pratama RP, Wibowo JW, Anto IAF (2017) Design of real-time weather monitoring system based on mobile application using automatic weather station. In: 2017 2nd international conference on automation, cognitive science, optics, micro electro-mechanical system, and information technology (ICACOMIT). IEEE, pp 44–47
10. Guo X, Song Y (2010) Design of automatic weather station based on GSM module. In: 2010 international conference on computer, mechatronics, control and electronic engineering, vol 5. IEEE
11. Kusriyanto M, Agusti AP (2018) Weather station design using IoT platform based on Arduino mega. In: 2018 international symposium on electronics and smart devices (ISESD). IEEE

A Smart and Intelligent Alcohol Detection System for Corporate Organization

Tejasvi Ghanshala, Vikas Tripathi, Prabhdeep Singh, and Bhasker Pant

Abstract In today's world, we find that life is becoming increasingly busy and hectic with each passing day, and as a result, employees of any corporate organization work extremely hard to meet project deadlines. Furthermore, as evidenced by numerous cases from the corporate sector, some of them used to drink after work and before work. So the issue is how to keep track of these activities in the office. To address such flaws in the system, a proposal has been made to detect people who come to work while inebriated. There will be no additional setup required in the office, according to the proposal. At each entry gate where each employee must punch before entering, a small alcohol sensor is all that is required. The alcohol sensor will detect each person's alcohol sensitivity and send the data to the server storage, where the database developers will perform the ETL process on the data and save it in the form of OLAP cubes, which will help in the future in generating reports with multidimensional data from which the admin and HR will get the record of each employee through application. In this way, the company can keep a hold on the employee, which will improve the company's rating and market growth.

Keywords Alcohol detection system · IoT · Cloud computing · Business intelligence

T. Ghanshala (✉) · V. Tripathi · P. Singh · B. Pant
Graphic Era Deemed to Be University, Dehradun, India
e-mail: tejasviuniversity18@gmail.com

P. Singh
e-mail: prabhdeepsingh.cse@geu.ac.in

B. Pant
e-mail: bhasker.pant@geu.ac.in

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_16

1 Introduction

Internet of things is not a new term now for anyone belonging to the technical field which was basically named by the professors at MIT nearly 20 odd years back [1]. In their invention, they stated that the things which we use in our daily routine life which could be anything like devices, microchips, sensors, smartphones are individually smart enough to be connected with each other so that the data can be collected and shared from any device with the help of another device placed anywhere. These devices when connected with each other through Internet can collect endless data which we can never think of or is out of our thinking range [2, 3].

So basically, this unique relation between the Internet and the things is all about the Internet of Things. Now imagine if in a house the refrigerators, Microwaves, Air conditioners, etc. has been connected with the Internet and then the data will be collected form them could be a huge saving to our time, money and brings a contribution to the safety of our house at the foremost [4, 5].

Some of us are working with corporate sector and we find that sometimes some of the employees or the housekeeping staff or the sweepers came to the office when they are even drunk which is an invitation to any misleading activity in the office, which at the end of the day spoils the reputation of the company just because of someone who is not bothered of it [6, 7]. As there are different towers in my office's building and each tower has multiple floors and each floor is subdivided into n numbers of ODC's so the system is well diversified into different sections and there is a punching machine fitted at each door which every employee has to punch for accessing the door but that punching machine can never detect that which employee is drunk so there is no restriction to such employee reaching his workplace or cabin. There is an admin department for each activity to manage things in the office's campus but for them as well it is not easy task to keep an eagles eye on this as the number of employees is quite large. So, there is proposal on this in such a way that a sensor will be deployed at each entry gate which will detect the alcohol and SMS module or Wi-Fi module will send the notification to the admin department that employee with particular employee id has entered in the office after being drunk. With this the admin department can take necessary actions on the employee which will restrict such cases in future and there will be no misleading for any male or female employee.

Internet of things (IoT) will help in collecting the data from the sensor which will be installed at the entry gates so that whenever each employee punches in at the punching machine his/her breath gets sensed by the alcohol sensor which is then connected to the Wi-Fi module with the help of Internet, through which the log of every employee gets to the admin department and if the person is drunk than the sensor will send the signal to the admin department that this particular employee with his/her employee code has come to the office despite of the fact that he was drunk so that the department could take the necessary action against that employee. In this way, there will be a great hold on the misleading activities which could happen in the office in the influence of alcohol which in result can thrash the company's reputation at the end of the day.

Business intelligence (BI) is an emerging technology in the today's world. BI is a vast domain which uses various data analyzing tools, data fetching tools, Data mining tools and on top of them the report building tools as well. The database developers have expanded their domain in BI by working on data mining, data cleaning, data warehousing, report building to support the strategic and tactical decisions made by the organization. Business intelligence is a cool and intelligent way to transform the raw and unmanaged data in a valuable form. The BI domain works in different phases, first comes the data integration in which three sub processes get into work and those are extraction of raw data or unmanaged data, transformation of that data to the specified location and then the loading of that data into the location upon which the further reporting will be done in further stages. Then comes the data warehousing stage where the various challenges get faced by the developers as to realize the ETL process and develop the OLAP cubes of various dimensions according to the type of data stored in the data warehouse. Once the data gets stored in the dimension format in cubes than comes the step to analyses that multidimensional data which is a typical part of this process as to read and understand data according to the clients need and requirement. When complete data analyzing is done by the developers the next part for them is to generate reports out of that multidimensional data stored in the cubes format.

2 Related Work

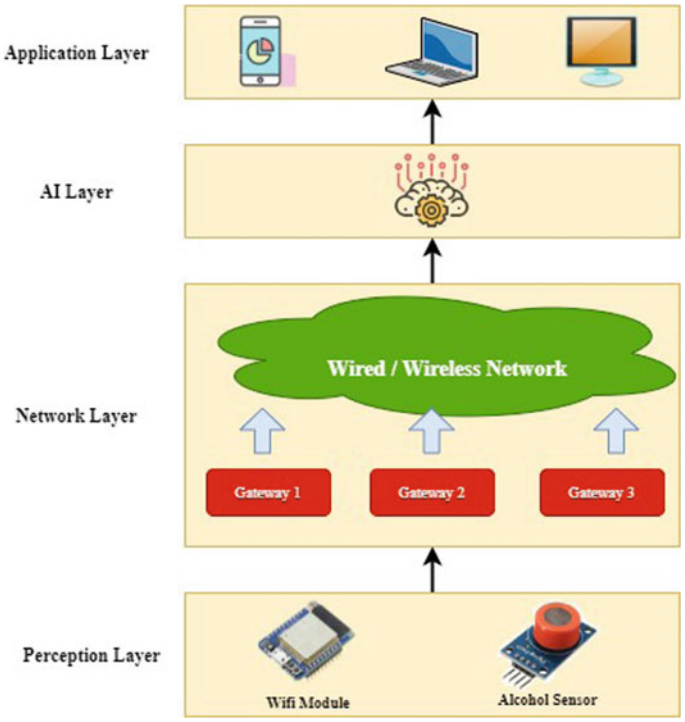
In the [8], the enhanced selective resonant frequency tuning approach was applied to detect methanol, ethanol, and 2-propanol in capacitive mode using a Pd/TiO_2 nanorod array/Ti device. [9] use a photoplethysmogram to collect the signal from an item in order to determine the advanced degree of alcohol consumption. In [10], a novel approach was used that is entirely based on Tunable Diode Laser Absorption Spectroscopy (TDLAS), a remote detection technology used to determine the alcohol content in a vehicle. In [11], the sensor supplies current linearly proportional to the molecules of alcohol from zero to extremely high concentrations, allowing the pic-microcontroller to verify the threshold value, which, when exceeded, automatically shuts down the relay and the automobile turns off. The objective of [12] is to design a system that captures the driver's iris picture by determining if the driver is intoxicated. The article will discuss hardware and software systems that will be used to create a Gabor filter method.

The system will include a CCD camera and an analog-to-digital converter that will be integrated into a MATLAB software to mimic the acquired picture, which will then generate a signal for the microcontroller and a relay circuit to manage the car's ignition. In [13], a portable Electronic nose (e-nose) chemical gas sensors have been shown to be an excellent tool for alcohol monitoring. In [14], the alcoholic concentration is determined using Non-Dispersive Infrared (NDIR), a technique that detects the degree of concentration without physical touch. The NDIR system operates by passing infrared light produced by an infrared source through the sample

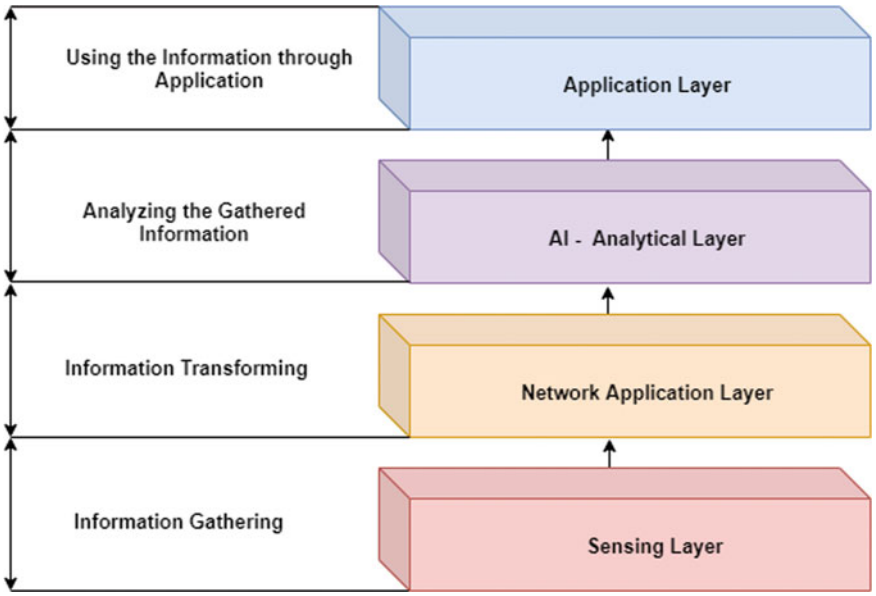
tube. Then, an infrared detector detects the infrared light that can be absorbed by the alcohol's hydrocarbon particles. The absorbed infrared light is then detected using a thermopile sensor. The result indicated that the optimal temperature for alcohol absorption is 60 °C, and that it is related to the magnitude of the concentration; the higher the concentration, the larger the absorbance value. The concept in [15] is to measure alcohol consumption levels using an alcohol detector attached to a printed circuit board. Which, in turn, disables the car's motor if the alcohol level exceeds the allowed limit. In [16], they present a framework in which they explain how they constructed the data warehouse from an existing information repository and how they managed to make intelligent report customizations based on the client's demands. In [17], the authors conducted an in-depth examination of how business intelligence technology works by emphasizing fundamental algorithms and recent scenarios. They also discussed the challenges and opportunities associated with BI in order for industrial informatics to connect with enterprise systems in the industry.

3 Proposed System

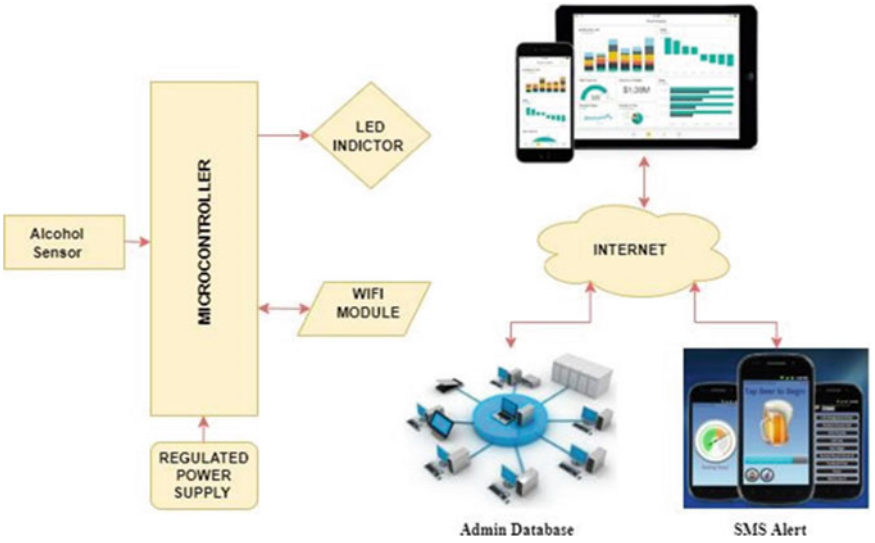
The proposed framework is shown in the above diagram that how the framework will work in the layered architecture when each layer will communicate through a proper channel. Coming to the Perception layer in which there will be two devices namely alcohol sensor and the Wi-Fi module. The alcohol sensor will detect the effected person's breath and sends the data to the next layer with the help of Wi-Fi module by communication with each other parallelly to the network layer through different gateways provided at the initial point of network layer which will be full equipped with the desired option of wired network using optical fiber or the wireless using 3G, 4G spectrum, LTE or 5G. Now this network layer will help the signal or data to travel distances through the medium which then traverses through different applications with the help of different middleware options available. When the data will be there in the applications than it will become easy for the admin department to fetch the data of the effected person on timely basis and based on the facts of such misconducts the effected person's appraisal or any other variables could be seized as the penalty of the misconduct in the office. To understand the whole concept lets dive into the layered architecture of the framework that how the system will work as shown in the below figure.



The above shown layered architecture of the proposed framework clearly shows that how the data will be captured, then transformed, analyzed and used for the eviction or penalizing the effected person in the organization.



Coming to the first layer from the bottom, in this layer there will be alcohol sensors installed at the entry gates of every tower through which each employee must enter while punching in his access card to the tower or building. While entering in the breath will be capture by the alcohol sensor which will be converted to the digital signals and the data will be sent across the network layer which will be equipped with wired or wireless technology (Wi-Fi, 3G, 4G, LTE, 5G). Through this layer the data will be transformed to the data warehouse or database and it will get saved over there for the analytical purposes. Once the data gets saved it will contain each employee's data whether he was drunk or not throughout the year and how many time he has come to office after drinking alcohol. Now third comes the BI—Analytical layer where the BI developers will fetch the data for the whole year and make it clean and valuable data so that they can get the number of employees who used to come office after drinking alcohol. After data mining, data cleaning, and analyzing the useful data will be saved at another location from where the applications or the end user can access the data for the evaluation purposes at the year end. When the cleaned data gets saved at a location then the end users or admin can see the data through their systems or the applications deployed for the same purpose. After making the reports of the data achieved from the Analytical layer the admin department and HR department could co-ordinate with each other for calculating the appraisal of the employee according to his behavior and attitude toward office in which the major concern will be coming to office after drinking alcohol. In this way, the proposed framework will help the organization in raising its standards and reputation in the market.



Algorithmic Explanation

Step 1: Connect the alcohol sensor's AOUT output to the Arduino's analog pin A0.
Step 2: Connect the alcohol sensor's DOUT output to the Arduino's digital pin D8.
Step 3: The LED's anode is connected to the Arduino's digital pin D13.
Step 4: Connect pin 4 to NODMCU GPIO 0.
Step 5: The LED's anode is connected to the Arduino's digital pin D13.
Step 6: When the LPG sense value is 12.
Step 7: Configure the baud rate.
Step 8: Connect the pin to the Arduino as an input.
Step 9: Assign the pin to the Arduino's output.
Step 10: Read the analog value from the AOUT pin on the alcohol sensor.
Step 11: Obtain the digital value from the DOUT pin on the alcohol sensor.
Step 12: Indicate the alcohol content.
Step 13: Indicate if the limit has been achieved as LOW or HIGH (above or underneath).
Step 14: If (limit = EXCESSIVE).
Step 15: A status indication LED illuminates.
Step 16: If the threshold is not achieved, the LED stays in the off position.
Step 17: Additionally, the computerized writer's alcohol sense will be LOW.

4 Conclusion

To conclude this paper, the best statement is that it contains an introduction to the smartest technology available, namely, IOT, which enables devices to communicate with one another over a network, business intelligence, which simplified data extraction, transformation, and loading, and labor savings, which reduce the cost of the entire concept. Additionally, after gathering the data, analysis was performed on it, resulting in reports from the multiple sources. If a business implements this solution on its campus, it will be the most effective gadget for detecting individuals who come to work inebriated, and it will be a cost-efficient option for any enterprise.

References

1. Singh P, Kaur R (2020) An integrated fog and artificial intelligence smart health framework to predict and prevent COVID-19. Glob Transitions 2:283–292
2. Singh P, Kaur R, Singh KD, Dhiman G (2021) A novel ensemble-based classifier for detecting the COVID-19 disease for infected patients. Inf Syst Frontiers 1–17
3. Gupta V, Gill HS, Singh P, Kaur R (2018) An energy efficient fog-cloud based architecture for healthcare. J Stat Manage Syst 21(4): 529–537
4. Singh P, Kaur R, Singh KD, Dhiman G, Soni M (2021) Fog-centric IoT based smart healthcare support service for monitoring and controlling an epidemic of Swine Flu virus. Inf Med Unlocked 100636
5. Singh G, Sharma S, Singh P (2013) Design and develop a honeypot for small scale organization. Int J Innov Technol Explor Eng (IJITEE) 2(3):170–174
6. Singh P, Kaur R, Dhiman G, Bojja GR (2021) BOSS: a new QoS aware blockchain assisted framework for secure and smart healthcare as a service. Expert Syst e12838
7. Singh P, Kaur R (2021) Implementation of the QoS framework using fog computing to predict COVID-19 disease at early stage. World J Eng
8. Dutta K, Bhowmik B, Bhattacharyya P (2017) Resonant frequency tuning technique for selective detection of alcohols by TiO_2 nanorod based capacitive device. IEEE Trans Nanotechnol
9. Izawa Y, Omura Y (2014) Reliable method to mask smoking effect on alcoholic intake detection based on photoplethysmogram signal analysis. IEEE Sens J 14:1418–1424
10. Shao J, Tang Q-J, Cheng C, Li Z-Y, Wu Y-X (2010) Remote detection of alcohol concentration in vehicle based on TDLAS. In: 2010 symposium on photonics and optoelectronic (SOPO)
11. James N, John TP et al (2014) Alcohol detection system. IJRCCT 3:59–64
12. Navarro LA, Diño MA, Joson E, Anacan R, Cruz RD (2016) Design of alcohol detection system for car users thru iris recognition pattern using wavelet transform. In: 2016 7th international conference on intelligent systems, modelling and simulation (ISMS)
13. Siyang S, Seesaard T, Lorwongtragool P, Kerdcharoen T (2013) E-nose based on metallotetraphenylporphyrin/SWNT-COOH for alcohol detection. In: 2013 IEEE international conference of electron devices and solid-state circuits (EDSSC)
14. Budiman F, Rivai M, Pambayun JPF (2016) Non-dispersive infrared (NDIR) sensor design and its application on alcohol detection. In: 2016 international seminar on intelligent technology and its applications (ISITIA)
15. Gupta, Ojha S, Kumar V, Singh V, Malav V, Gramothan R (2016) Alcohol detection with vehicle controlling. Int J Eng Manage Res 6

16. Keerin P (2016) Development of business intelligence solution for personnel administration. In: 2016 second Asian conference on defence technology (ACDT)
17. Duan L, Da Xu L (2012) Business intelligence for enterprise systems: a survey. IEEE Trans Industr Inf 8:679–687

Legal Frameworks and Issues of Social Media Use for Politics

Getachew Hailemariam Mengesha, Elefelious Getachew Belay, and Moges Ayele Asale

Abstract This study intends to explore legal frameworks and ethical considerations that would regulate social media use to promote politics and for social interaction in the context of Ethiopia. Apparently, social medias are found convenient, effective, and less expensive to market political agendas and to reach out voters easily. Particularly, they are ideal to attract young voters who have low interest in politics because of ease of generating contents, commenting and reflecting on posts, ease of sharing contents to large audience breaking down barriers to freedom of expression. Data for the study were gathered through domain expert involved focus group discussion (FGD). Thematic analysis approach was used to generate themes and insights from the data. The study reveals interesting results on two key relevant issues: media law and individual rights. From the perspectives of media law and policy, the study uncovered paucity and loopholes connected to nurturing a culture of constructive critique on political matters, the need to have sound communication policy at all key governmental institutions, weak regulatory system in terms of crafting and enforcing relevant media laws, the tendency of demeaning accountability on the side of social media companies and social media users, the need to explore legal and diplomatic pressures on social media companies to discharge their corporate social responsibilities by protecting the platform they have developed not to be used to incite conflict, violence, and human right abuses. With regard to upholding fundamental human rights, social medias have provided great opportunities to the general public to easily air out their views and promote political agendas they support. Social medias also paved the way to realize freedom of speech, which presupposes orderly liberty that

G. H. Mengesha (✉)
School of Information Science, Addis Ababa University, Addis Ababa, Ethiopia
e-mail: getachew.hailemariam@aau.edu.et

E. G. Belay
School of Information Technology and Engineering, Addis Ababa Institute of Technology, Addis Ababa University, Addis Ababa, Ethiopia
e-mail: elefelious.getachew@aau.edu.et

M. A. Asale
School of Psychology, Addis Ababa University, Addis Ababa, Ethiopia
e-mail: moges.ayeleA@aau.edu.et

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_17

do not grant a license to threaten individual's existential rights opening the door for gross human right violations.

Keywords Social media · Media laws · Communication policy · Human right · Existential rights · Freedom of speech

1 Introduction

The advent of social media (SM) has brought major social transformations throughout the world. Social media tends to disrupt the conventional mainstream mass media. These traditional medias are lagging behind in terms of providing instant breaking news as compared to the social media sites. Rather, they are widely seen echoing events already broadcasted by the social media. The wider reach of social media fostered social activism and freelance journalism. True and fake messages are equally circulating through social media. Furthermore, social media is becoming a powerful tool for coordinating social and political movements around the world. For instance, Twitter and other social media outlets are extensively used to call for virtual and physical protests. In this regard, the London 2011 youth protest, the Iranian protest against Ahmadinejad in 2009, the Arab Spring that outbreak in North Africa and swiftly spread through the Middle East can be cited as prominent cases. Further, the social movement in Ethiopia that culminated with toppling down the EPDRF government in 2018 was partly orchestrated by activists that used the social media extensively. These incidents clearly indicate that social media sites are not simple communication tools, instead they are standing out as tools that create new social realities.

While what transpired so far are positive impacts, social media tools are also being used to incite conflicts and instability in many places around the world. For instance, Belay et al. [3] state that social media tools are used to instigate violence and genocide across many regions in Ethiopia. This has raised a heated debate surrounding the corporate social responsibilities stance of social media companies like Facebook, the dilemma between freedom of expression and existential rights of individuals. Quite recently, social media companies exercised unprecedented power by removing message posted on authorized sites of heads of state such as Donald Trump and Abiy Ahmed and moved to the extent of terminating pages of various influential activists around the world. This act has led to a mixed view spanning from unfair meddling of social media companies in political affairs to courageous act that demonstrates even heads of states need to respect the rights of social media user communities. The unprecedented powerfulness of social media companies coupled with the massive capability of social media technologies in shaping social reality, calls for a regulating mechanism before the world encounter new form of monopolistic traps that control every aspects of human life.

This study aims to explore legal frameworks and ethical considerations that would regulate social media use while promoting political campaigns and contests. Apparently, social media sites are found convenient, effective, and less expensive to market political agendas and to easily reach out voters. Particularly, they are ideal to attract young voters who have low interest to politics because of ease of generating contents, sharing, and commenting on posts. Astuti and Hangsing [1] noted social media sites enable to reach out large audiences overcoming geographical barriers and shortening the communication channel. The authors further note that multimedia features of social media makes easier to present political programs to the general public in order to recruit and convenience voters.

Regarding use of social media, a polarized views are observed ranging from strict control to unconditional free use of the technology as espoused in basic human rights or freedom of expression. Nevertheless, this study posits that a delicate balance between freedom of expressions and existential rights of individuals must be maintained. Furthermore, most social media technologies are owned by companies in a developed world, and they are abided by their home countries legislation and code of conducts. In the absence of sound cross-border governance frameworks, developing nations like Ethiopia tend to be disproportionately affected by inappropriate use of social media platforms. Perhaps, in pursuit of maintaining internal peace and security, developing countries need to limit social media platforms with which they have signed cross boarder governance agreements. However, in many developing countries, there is a paucity of strong social media legislation. Coupled with absence of global social media governance framework, the adverse effect of social media platforms is devastating many countries around the world, particularly African countries. Further, as [15] Michailova and Hutchings (2006) state developing nations trail with lack of a well defined legal systems that make it difficult to control inappropriate uses such as dissemination of false narratives to flare ethnic conflicts. Particularly, in most African countries, an individual tends to have two identities: individual and collective. Some morally disengaged irresponsible individuals deliberately use these social media to provoke civil unrest and to cause physical and psychological damages. Unless, nations like Ethiopia are bold enough either to set up their homegrown social media platforms or seek mechanisms to regulate the use of the existing popular global social media platforms, their internal stability, and security will fall at stake.

The remaining part of the paper is structured as follows. First, the methodology used for the study is described. Second, key findings of the study will be presented. Third, discussion and reflection on the study are provided. Finally, the paper ends with conclusion.

2 The Study

Data for the study were collected through focus group discussion (FGD) conducted as a panel discussion on December 24, 2020. The panel session was fully recorded (voice and video). The panel discussion was moderated by media professionals and later

broadcasted by various media outlets. Five panelists drawn from social psychology, law, digital media, cybersecurity, and journalism took part in the panel discussion. Predesigned interview protocol was used to pose questions to the panelists in accordance with their professional background. Later, the panel recordings were transcribed. The data were analyzed using thematic analysis approach. The research teams (four team members) first conducted the coding independently and presented their codes to the group in organized discussion sessions. After an in-depth discussion, 84 initial codes were identified. The research team further refined the codes through a series of discussion sessions and eventually about 17 themes emerged, of which ten were high-level abstractions each containing at least three sub-themes. This paper deals with legal issues emerged as a prominent theme.

3 Findings

This section enlists key themes emerged surrounding legal and basic human right issues connected to use of social media platform in Ethiopian context. Table 1 shows key constructs emerged from media policy and individual rights perspectives.

As indicated in Table 1, pertinent themes have emerged in connection to media policy and individual right. With regard to media policy and law, first, respondents of the study underscored on the commitment of the government to promote a culture of healthy political dialog posing constructive critique on governments policy without fear of prosecution. In the short run, this might cause unease in the socio-cultural landscape, where power distance as posited by Hofstede [10] is an enduring norm in Ethiopian society. However, it will have a profound long-term impact that nurture a culture of constructive critique and debates on political matters. In this regard, the panelist stated that

> Striking balance between free speech and national security is difficult. For instance, in USA, the law that forbids inciting protest and violence was passed in 1919, over 100 years ago. According to this law, if a speech clearly leads to protest and violence, it will be treated as criminal. However, the law was revised in 1969 and declares that if a speech has clear and immediate violent effect, it is considered criminal. Although the law favors free speech,

Table 1 Media policy and individual right

No	Media law and policy	Individual right
1	Nurturing a culture of constructive critique	Opportunity to air out views
2	Communication policy	Freedom of speech and protest
3	Weak regulatory system	Orderly liberty
4	Undermining accountability	Existential right
5	Legal and diplomatic pressure	Violation of human right

> about 2000 people have been arrested in the country due to offense related to their speeches. Further, the 1948 Genocide Convention strictly forbids inciting genocide in speech and deeds.

Second, a part from grand media use violations being made by dozens of social activists, opposition politicians, and the general public, and some government officials were seen making dangerous speeches that incite conflicts among the society. Many government officials tend to fall in untold dilemma and role conflicts while running federal offices they are appointed for. They were seen tensioned failing to balance national integrity and ethnic interests emanating from their constituency and ethnic origins. As a remedy to this, respondents suggested that

> Government officials need to be abided by a communication policy, which is not yet in place. The messages disseminated by a government official are very sensitive, and therefore, they need to be cautions of their speech while appearing to conventional mass media as well while using social media.

Third, the government needs to seek legal and technical mechanisms to regulate leading social media platforms. Apparently, many developing nations have a weak regulatory system. In this regard, respondents cited the following as an example:

> If we consider the case of Myanmar, Facebook played major role in the ethnic cleansing of the Rohingyas. The technology companies have got a big financial power and huge economic interest, running multi-billion-dollar business. Hence, governments need to demand for internal control based on their objective conditions; they have to adopt a customized usage of social media.

Fourth, conventional wisdom and societal expectations and normative laws denote that individuals and entities are held responsible for their deeds and speeches regardless of the media they have used. The fact on the ground reveals that some individuals seem to avoid responsibilities for the content they post and disseminate via social medias. In this regard, respondents note that

> Some government officials use inflammatory languages contrary to the prevailing social values; we do not see them assuming responsibility for their speeches. Therefore, government officials need to avoid information and speeches that tend to incite discontent and conflicts. Similarly, mainstream media journalists need to undertake fact and public-sensitivity checks before citing news obtained from social media sources.

Fifth, legal and diplomatic pressures are needed on major social media companies such as Facebook and Twitter to discharge corporate social responsibilities. The platform they have developed and managing should not be used to carry out genocide and human right abuses. Reflecting on this fact, respondents state that

> Social media platforms need to be held accountable for any damage that may occur due to information disseminated by their users. They ought to accept that their business promotion platforms should not be sources of chaos and danger to others. For instance, Facebook officials acknowledged the danger that the platform caused in Burma. The same thing is happening in Ethiopia; if there is a danger of genocide due to the use of Facebook, the company has to be held accountable.

In connection to upholding and violating basic human rights, the study generated varied insights. First, social media outlets provided opportunities to millions around the world to air out their views freely and attain psychological relives. Ideas and feelings of voiceless people would remain unheard, if convenient social media outlets did not exist. Furthermore, social media outlets made possible freedom of speech and the right to protest unjust practices. Inline to this, respondents state that "if citizens get a chance to articulate their views, they would have less desired to get involved in protests."

In nations, where democratic culture has not developed, people tend to misconceive democratic freedoms and the associated responsibility. Reflecting on this issue, respondents state as follows:

> Freedom of expression does not mean unlimited freedom and lawlessness. In any country, democracy does not mean lawlessness and unregulated act; democracy is an orderly liberty. Hence, the public need to understand these facts and act accordingly.

The study also revealed a mixed view of government agencies and the general public in relation to exercising freedom of expression and respecting existential right of individuals. Evidently, existential right of individuals overrides freedom of expressions and other rights. Surrounding this issue respondent's state, the following:

> We have to know that people's right to stay alive is far more important than freedom of expression. The prevailing tendency of demeaning existential rights of individuals in the name of collective right or ethnic right is a clear violation of human right.

Finally, gross human right violations are observed as a result of in appropriate use of social media that flared ethnic tensions. While many individuals commit mistakes unintentional in the course of interacting through social media sites, there are a handful of individuals with political and religious intentions deliberately produce and disseminate dangerous information, which incited conflicts, displacement, and killings of on specific ethnic groups. For instance, respondents noted the following:

> Especially in our country, many people are sensitive to ethnic-based information, and some individuals take such advantages. They create fake personalities with different ethnic affiliations. They purposefully create hostile dialogs in which Mr. X insults Mr. Y so that the public perceive as if there was a conflict between two ethnic groups. Hence, they artificially create a condition in which a real people of the two communities enter into conflicts on the ground.

4 Discussion and Reflection

Undoubtedly, social media has transformed the way people have been interacting overcoming barriers of distance and time. This study explored the positive impact as well and the danger aroused by improper use of social media platforms. In this section, we reflect on various pressing issues related to Ethiopian cyber law and regulations; the quest to balancing freedom of speech, national security, and citizens existential rights; conceptions surrounding freedom of speech, democratic culture, and use of

social media; human rights violations and social media use; social media monitoring and legal concerns and dealing with the dilemma between open vs controlled use of social media services.

Ethiopia does not have a comprehensive legal instrument which regulates privacy and data protection. However, there are rules that directly or indirectly deal with data privacy and protection. In addition, there are an ongoing effort to draft legislations by government agencies like Ministry of Innovation and Technology (MInT) and Information Network Security Agency (INSA). Cybercrime-related legislations received due attention in 2004 when the Penal Code of Ethiopia enacted in 1957 was revised. Halefom [9] states that the first legislative statement concerning cybercrime appeared in the 2004 Criminal Code of Ethiopia. Later, Ethiopian National Information Security Policy document was developed in 2011. The policy considers cybersecurity as a critical national security and human right issue. This document served as a basis when INSA was revitalized and institutionalized by Regulation No. 250/2011 of Council of Ministers and Proclamation No. 808/2013. Furthermore, the House of Representatives of Ethiopia passed another Computer Crime Proclamation No. 958/2016. The proclamation addresses broader issues that involve information privacy, security, dissemination of offensive contents, and grants power to law enforcement bodies to undertake surveillance as the need arises. In general, the constitution of the country, the National Information Policy drafted in 2011, Media Proclamation No. 560/2008, and Computer Crime Proclamation 958/2016 are the existing legislations to deal with cases related to social media abuses.

Many nations around the world seem to have fallen in a dilemma with regard to balancing free use of social media and ensuring national security and preserving social and religious norms. This is evident because of the fact that regardless of the nature of the regime: democratic or authoritarian, social media, and Internet disruptions widely observed when the regime is challenged [12]. These authors further noted that modern communication technologies including social media provided unprecedented opportunities to easily establish extensive ties and create social capital to mobilize political actions. Apparently, digital communications have radically transformed the way political campaigns have been conducted for a longer period of time [12, 5, 13, 18]. Technology savvy political parties outperform strong opponents through strategic use of social media. This fact tends to compel many states to temporarily disrupt to social medias and the Internet until they overcome tense civil unrest which would be further flared by social media [12]. On the other hand, some activists with large number of social media followers fail to discharge their responsibility of imposing self-censorship and discipline while posting and disseminating contents on social media. In the name of freedom of expression, disseminating sensitive ethnic and religion related content tend to cause social unrest in Ethiopia.

Access to basic information and freedom of speech are considered as a fundamental human right [19]. However, these rights will not be automatically ensured without independent and professional medias and without empowering citizens to build capacity to engage in a public debate [19]. Coupled with the national law, the UNESCO declaration regarding freedom of speech can serve as a guiding tool while investigating freedom of speech cases in relation to social media use. Freedom

House (2019) highlighted how government and non-government actors inappropriately use social media to violate human right and to undertake election interference. Abuse of these medias tend to take place using technical measures to halt connectivity, through dissemination of false information and hacking social media accounts. Further, governments tend to engage in applying laws that bans or instance applying the law that bans the defamation of leaders or rulers (Shahbaz and Funk 201, cited in [2]).

Governments in many parts of the world alleged disinformation and misinformation to justify laws that restricts social media use. For instance, human rights watchdog condemned that from June 2017 to May 2018 about seventeen countries enacted laws to limit social media in the name of controlling fake news [2]. Although, state intervention in Internet use related affairs first occurred in Germany, in the period ranging from 1995 to 2011 about 566 cyber control scenarios that involve 101 countries has occurred [12]. Further, the authors state that national security and the need to counter terrorism are the most cited reasons for states intrusion and surveillance of network system including various mechanisms such as shutting down Websites, arresting bloggers, activists, and journalists even in the worst-case scenario disrupting the Internet at all for long or short durations.

Social media is now associated to basic human rights and freedom of speech and regarded as an instrument to nurture democratic culture. Social media tools are convenient to easily express opinion, to reflect on other people's content and to convey message to large group of audiences quickly and at a least cost. These features of social media removed barriers and helped to arose the taste of the youth to politics and have ensured their active and influential participation in political matters [1].

Recent unfolding events indicate that social media abuse is causing human rights violations and threatening existential rights of individuals. Social media is not just a place where crime and justice are discussed; it is a place where crime happens, and the quest for justice is increasing rapidly. Reflecting on contentious issues surrounding social media use, crime, and justice concerns, Salter [16] states that social media is host to debate and engage in conflict over what constitutes crime and justice. He further notes that this disagreement is evident not only between users but also between users and social media companies.

Crime and justice on social media have been analyzed within a critical theory perspective, recognizing that online abuse and justice-seeking take shape within the communicative potentialities offered by social media [16]. Several unresolved issues exist surrounding social media uses. Still, there is an ongoing debate whether social media use monitoring measures are legally sound and yet untold dilemma arouse in connection to open versus controlled use of social media services.

Technological advancements can improve the economy of a given society. However, the misuse of social media has often had negative impacts on certain individuals and groups. There are some damaging elements in social media use such as spreading false news or hoaxes, the utterance of hatred, the use of ethnic-based stereotyping, memes about state equipment that are unethical, even acts of treason we often encounter in social media. Hence, law enforcement must be truly active and observant in responding to improper and illegal actions [17].

5 Conclusion

This study affirms that social media abuse is leading human rights violations and is threatening existential rights of individuals in Ethiopia. Social media is becoming a place where crime happens. Furthermore, social media sites are not simple communication tools, rather they turn out as instrument that creates new social realities. Modern social media backed by artificial intelligence (AI) algorithms have started shaping opinions and influencing vote outcomes. For instance, Twitter bots influenced votes in favor of a particular candidate shaping voters opinion generating some information and building a sense of consensus in the society [14, 6, 7]. Therefore, it is high time to protect democratic freedom and human rights by ensuring that social media and associated Internet technologies do not become instrument of repression and human right violation. Several countries approved laws that would restrict in the name of thwarting fake news, misinformation, and disinformation. However, institutions like human rights watchdogs are criticizing this act as a violation of freedom of expression. The world is grabbling with several unresolved issues surrounding social media uses. Still, there is an ongoing debate whether social media use and monitoring measures are legally sound. Further, an untold dilemma arouses in connection to open versus controlled use of social media services. Moreover, big social media companies are gaining unprecedented power and seen failing to discharge their corporate social responsibilities as intended. Developing nations like Ethiopia are at a crossroad lacking resolutions whether to let their citizens to freely use global social medias or impose restrictive measures or whether they need to seek technical solutions to develop homegrown social media technologies and craft relevant legislations that shape the use and governance of these technologies. These are perplexing issues that demand legal, technical, ethical, and governance mechanisms to ensure sound use of social media for social good.

References

1. Astuti PAA, Hangsing P (2016) Social media and effort to coax young voters in Jakarta, Indonesia. Scriptura 6(1):1–13
2. Aswad EM (2020) In a world of “Fake News”, What’s a social media platform to do? Utah L Rev 1009
3. Belay EG, Mengesha GH, Asale MA (2020, July) The tributes and perils of social media use practices in ethiopian socio-political landscape. In: International conference on human-computer interaction. Springer, Cham, pp 199–209
4. Berger G (2009) Freedom of expression, access to information and empowerment of people
5. Bimber B (2005) Reconceptualizing collective action in the contemporary media environment. Commun Theor 15:365–388
6. Gagliardone I, Pohjonen M (2016) Engaging in polarized society: social media and political discourse in Ethiopia. In: Digital activism in the social media era. Palgrave Macmillan, Cham, pp 25–44
7. Gorodnichenko Y, Pham T, Talavera O (2021) Social media, sentiment and public opinions: evidence from #Brexit and #USElection. Eur Econ Rev 103772

8. Hailu H (2015) The state of cybercrime governance in Ethiopia. Article published on ResearchGate. Available at https://www.researchgate.net/publication/322234805_THE_STATE_OF_CYBERCRIME_GOVERNANCE_IN_ETHIOPIA.
9. Halefom H (2015) The state of cybercrime governance in Ethiopia. Abyssinia Law. https://www.abyssinialaw.com
10. Hofstede G (1983) National cultures in four dimensions: a research-based theory of cultural differences among nations. Int Stud Manage Organ 13(1-2):46–74
11. Howard PN, Agarwal SD, Hussain MM (2011) When do states disconnect their digital networks? Regime responses to the political uses of social media. Commun Rev 14(3):216–232
12. Howard P, Sheetal A, Muzammil H (2010) When do states disconnect their digital networks? Regime responses to the political uses of social media. The commun Rev 14:1
13. Howard PN (2010) Digital origins of dictatorship and democracy: the Internet and political Islam, New York, NY: Oxford University Press
14. Hänska M, Bauchowitz S (2017) Tweeting for Brexit: how social media influenced the referendum
15. Michailova S, Hutchings K (2006) National cultural influences on knowledge sharing: a comparison of China and Russia. J Manage Stud 43(3):383–405. https://doi.org/10.1111/j.1467-6486.2006.00595.x
16. Salter M (2017) Crime, justice and social media. Routledge
17. Situmeang DM, Siringoringo M (2018) Social media abuse that defends particular parties reviewed from law of ITE number 11 in 2008 concerning electronic information and transaction. Int J English Literature Social Sci 3(5):265270
18. Still, B. (2005). Hacking for a cause. First Monday 10:1. http://irstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1274/1194. https://irstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1274/1194
19. UNESCO (2009) Freedom of expression, Access to information and empowerment of people. https://www.unesco.org/en

Prevalence and Determinants of Mobile Health Applications Use Among Saudi Adults

Hassan Kasim Haridi, Saad Alsaleh, Sulaiman Alzabin, Mohammed Almasabi, Abbas Almakrami, Ali Al-Swedan, and Abdelaziz Aman

Abstract Background: mHealth applications have added new dimensions for managing individual and community health and healthcare aspects. To date, information about the prevalence and factors associated with mHealth applications' use among the Saudi population is scarce. Objectives: This study aimed to estimate the prevalence and to explore determinants of mHealth Apps use among Saudi adults. Methods: A cross-sectional study was carried out to approach patients and caregivers using healthcare facilities in Hail region, Saudi Arabia. Results: Overall, 470 participants completed the survey questionnaire. Females were 336 (71.6%) and 288 (61.3%) were university educated. Almost all 464 (98.7%) participants were smartphone owners and 423 (90.0%) of them have Internet access, however, 268 (57.0%) only were mHealth applications users. The multivariate logistic regression model revealed that higher education ($P = 0.023$), easy to access and upload the applications ($P = 0.030$), having no difficulty in registration ($P = 0.005$), little effort needed for mastering the applications (0.005) were the independent factors favoring mobile health applications use. Contrariwise, feeling discomfort in using the applications ($P = 0.003$) and having concerns regarding violation of privacy ($P = 0.036$), were factors independently associated with not using these applications. Conclusion: While almost all the participants were smartphone owners and have Internet access, only a fair percent use mobile health-related apps.

Keywords mHealth apps · Prevalence · Determinants

1 Introduction

The use of mobile and wireless technology to help people reach their health goals (mHealth) has the prospective to change health care delivery around the world. This

H. K. Haridi (✉) · M. Almasabi · A. Almakrami · A. Al-Swedan · A. Aman
General Directorate of Health Affairs, Najran, Saudi Arabia
e-mail: hassankasim@hotmail.com

S. Alsaleh · S. Alzabin
General Directorate of Health Affairs, Hail, Saudi Arabia

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_18

shift is being fueled by a potent confluence of variables. The fast developments in mobile technologies and applications and expansion of mobile cellular network coverage raised new opportunities for incorporating mobile health into the available eHealth services [1].

The 2030 Agenda for Sustainable Development emphasizes that the growth of information and communications technology, as well as global interconnection, has enormous potential to speed up human progress, close the digital divide, and establish knowledge societies [2].

Patient eHealth tools are a wide variety of patient-focused applications and services, like mHealth applications, personal monitoring devices, and Internet resources. Patients, family, and careers are all possible target users [1]. In terms of promotive, preventative, and curative aspects of healthcare, mobile technology and mHealth applications have contributed new dimensions. Healthcare workers are increasingly embracing mHealth applications to support their patients in making good lifestyle changes and empower them as well as healthy people to promote their health. These applications may also be useful in large gatherings of Muslims in Saudi Arabia during Hajj and Umrah to identify health hazards, prevent and manage health problems, and provide the health authority with a good opportunity to monitor, analyze, and develop health plans for a safer Hajj, Umrah, and visits.

Exploring the use of mHealth applications and factors influencing their use is critical to gain a better understanding for better utilization. We aimed in this research to assess the prevalence and determinants of mHealth applications use among healthcare users in Saudi Arabia.

2 Subjects and Methods

Our study was conducted using a cross-sectional survey design. Adult patients, caregivers, and visitors to healthcare services in the Hail province, northern Saudi Arabia were the target of this survey. Adult literate males and females aged 18 and up were approached. An anonymous, structured questionnaire was used to collect data. There were 24 questions in total, divided into five sections. The first section included 5 statements to address the perceived benefits of mHealth applications. The second section consisted of 4 questions aimed to determine how easy it to utilize these applications. The third section consisted of five questions designed to elicit concerns and obstacles for the application's use. Two questions were included in the fourth section to know about the participants' expectations and needs in terms of mHealth applications. The participants' sociodemographic factors (age, gender, education), as well as their use of technology, were assigned to the last section (having a cellular phone, internet access, and currently or previously using mobile health apps).

Participants' responses to statements about the perceived benefits, ease of use, concerns and barriers, and perceived expectations in response to their needs from mHealth applications were measured using a five-point Likert scale, ranging from strongly disagree (one point) to strongly agree (5 points).

The participants of the study were recruited using a two-stage random sampling procedure. At first, simple random sampling was used to define 4 healthcare facilities representing primary, secondary, and tertiary health care levels. Two hospitals and two primary health care centers were randomly selected. Systematic random sampling technique was employed in the second step to include the study participants, who were either patients, caregivers accompanying the patients, or visitors to the health setting. The sample was divided evenly between the participated health care settings.

The study sample was estimated allowing for the largest possible sample size. We assumed a 50% respondents' response for the main outcome of interest (mHealth applications use) at a 95% confidence level and a 5% margin of error. A minimum sample size of 384 was considered using the Krejcie and Morgan sample size table.

Our study was conducted during the first quarter of 2019. To ensure adequate participation, data collectors distributed 600 questionnaires to the eligible participants.

The Regional Bioethics Committee of the General Directorate of Health Affairs, Hail region, Saudi Arabia approved the study protocol and data collection tool on October 24, 2018. The voluntary participation and anonymity of the participants' identities and confidentiality of their responses were assured.

The statistical package of the Epi Info program version 7 was used to evaluate the subjects' demographic data for descriptive statistics. The frequency means, and standard deviation of the study variables was expressed, within a 95% CI. The Chi-square test was employed to decide the significant statistical difference in proportions. We carried out multivariate logistic regression analysis to identify independent variables associated with using mobile health applications among participants. A two-sided p-value ≥ 0.05 for all tests indicates a statistically significant difference. The instrument's Cronbach's alpha was (0.84), indicating adequate reliability.

3 Results

Overall, 470 questionnaires were returned complete out of 600 questionnaires distributed, giving a 78.3% response rate. Table 1, depicts the participants' characteristics.

Table 2, demonstrates the perceived benefits, ease of use, concerns, barriers, and expectations about mHealth applications among healthcare service users and non-users who have cellular phones and Internet access.

Table 3, depicts the result of the multivariate logistic regression model capturing independent factors linked to using mHealth applications among the study participants. Higher education (OR, 2.26; $P = 0.023$), easy to access and upload the applications (OR, 2.14; $P = 0.030$), having no difficulty in registration (OR, 2.92; $P = 0.005$), little effort needed for mastering the applications (OR, 2.39; $P = 0.005$) found to be independent factors favoring using mHealth applications. Contrariwise, feeling discomfort in using the mHealth applications (OR, 0.32; $P = 0.003$) and

Table 1 Participants' characteristics ($n = 470$)

Participants	*n*	%
Age in years		
<20	35	7.4
20–29	169	36.0
30–39	158	33.6
≥40	108	23.0
Mean (SD)	32.44 (10.46)	
Gender		
Male	133	28.4
Female	336	71.6
Education		
Elementary	59	12.6
High schooling	123	26.2
University/higher	288	61.3
Access to and utilization of mHealth applications		
Owns a mobile telephone	464	98.7
Have access to the internet	423	90.0
Ever used mHealth applications	268	57.0

Abbreviations: mHealth apps, mobile health applications

having concerns regarding the violation of privacy (OR, 0.53; $P = 0.036$), were factors independently associated with the non-use of mHealth applications.

4 Discussion

Almost all participants in our study (98.75) have smartphones, with more than ninety percent having Internet access, which indicates that Saudi society has a good potential to reach a high level of information technology utilization. However, our survey indicates that a still fair percentage of people (57.0%) who utilize mobile health applications, which implies the need to raise the awareness of the people in the community toward the benefits of such programs, and develop need-based trusted affordable programs with good users support.

In our study, the prevalence of mHealth applications usage was greater than a 22% estimate in a US national trends survey in 2015 [3] and 11.7% found in the prior 2012 US Survey [4]. This increasing prevalence can be explained by the recent timing of our survey, the rapid adoption of mHealth applications by people all over the world, as well as the rapid development of new applications tailored to people's health needs.

Table 2 Perceived benefits, ease of use, concerns, barriers, and expectations about mobile health applications (mHealth apps) among the participants who have mobile phones and Internet connections ($n = 423$)

Elements	mHealth apps users $n = 249$; 58.9%	mHealth apps non-users $n = 174$, 41.1%	*P-value*
Perceived benefits/usefulness			
Facilitate registration and health file update			
– Agree	233 (60.2)	154 (39.8)	0.046
– Uncertain/disagree	15 (42.9)	20 (57.1)	
Reserve time for booking appointments			
– Agree	219 (61.0)	140 (39.0)	0.034
– Uncertain/disagree	30 (46.9)	34 (53.1)	
Facilitate health service provision			
– Agree	215 (62.0)	132 (38.0)	0.086
– Uncertain/disagree	34 (50.7)	33 (49.3)	
Ease medical report request			
– Agree	232 (58.9)	162 (41.1)	0.978
– Uncertain/disagree	17 (58.6)	12 (41.4)	
Improve the quality of service			
– Agree	217 (60.1)	144 (39.9)	0.209
– Uncertain/disagree	32 (51.6)	30 (48.4)	
Perceived ease of use			
It is easy to access and upload mHealth apps			
– Agree	224 (64.6)	123 (35.4)	<0.001
– Uncertain/disagree	25 (37.3)	42 (62.7)	
I have no difficulty in using mHealth			
– Agree	196 (65.8)	102 (34.2)	<0.001
– Uncertain/disagree	53 (45.7)	63 (54.3)	
Little effort needed for grasping mHealth apps			
– Agree	194 (65.8)	101 (34.2)	<0.001
– Uncertain/disagree	55 (46.2)	64 (53.8)	
I face technical problems with mHealth apps			
– Agree	178 (62.9)	105 (37.1)	0.093
– Uncertain/disagree	71 (54.2)	60 (45.8)	

(continued)

Table 2 (continued)

Elements	mHealth apps users $n = 249$; 58.9%	mHealth apps non-users $n = 174$, 41.1%	*P-value*
Perceived concerns			
I have concerns regarding violating my privacy when using mHealth apps			
– Agree	56 (48.3)	60 (51.7)	0.006
– Uncertain/disagree	193 (63.1)	113 (36.9)	
I have concerns regarding misusing my personal data when using mHealth apps			
– Agree	75 (51.4)	71 (48.6)	0.023
– Uncertain/disagree	174 (62.8)	103 (37.2)	
Perceived barriers			
I feel difficulty in registration and dealing with mHealth apps			
– Agree	63 (55.3)	51 (44.7)	0.365
– Uncertain/disagree	186 (60.2)	123 (39.8)	
I need the help of others in using mHealth apps			
– Agree	73 (48.3)	78 (51.7)	0.001
– Uncertain/disagree	176 (64.7)	96 (35.3)	
I feel discomfort in using mHealth apps			
– Agree	39 (43.8)	50 (56.2)	0.001
– Uncertain/disagree	210 (62.9)	124 (37.1)	
Perceived expectations			
There is a need to have a unified health file to be accessed at any health facility countrywide			
– Agree	241 (60.1)	160 (39.9)	0.028
– Uncertain/disagree	8 (36.4)	14 (63.6)	
A need for health risk assessment application			
– Agree	231 (59.7)	156 (40.3)	0.258
– Uncertain/disagree	18 (50.0)	18 (50.0)	

Table 3 The fitted model* of multivariate logistic regression analysis capturing independent variables associated with using mobile health applications (mHealth applications) among healthcare service users

Factors†	aOR	95% CI	*P*-value
Education			
– Elementary	1		
– High school	1.3499	0.6335–2.8764	0.4369
– University	2.2611	1.1171–4.5764	0.0233
It is easy to access and upload mHealth apps			
– Agree	2.1365	1.0770–4.2384	0.0298
– Uncertain/disagree	1		
I have no difficulty in registration and dealing with mHealth apps			
– Agree	2.9235	1.3811–6.1887	0.0051
– Uncertain/disagree	1		
Little effort needed to master mHealth apps			
– Agree	2.3919	1.3012–4.3967	0.0050
– Uncertain/disagree	1		
Frequently, I face technical problems with mHealth apps			
– Agree	0.5636	0.3007–1.0563	0.0736
– Uncertain/disagree	1		
I feel discomfort in using mHealth apps			
Agree	0.3232	0.1513–0.6905	0.0035
– Uncertain/disagree	1		
I have concerns about violating my privacy			
– Agree	0.5337	0.2970–0.9590	0.0357
– Uncertain/disagree	1		

Abbreviations: mHealth apps, mobile health-related applications; aOR, adjusted odds ratio.
†Demographic variables in Table 1 and statistically significant variables in Table 2 were contained in the multiple logistic regression analysis.
A p-value ≤ 0.05 indicates a statistically significant difference.
OR > 1 indicates a positive association, while OR < 1 indicates a negative with the outcome.
*The fitted model log-likelihood = 464.71.

Our study revealed that only education had an independent effect on the use of mHealth applications among demographic variables (age, gender, and education). Participants with higher education in our study were significantly more probable to use mHealth applications than others with lower education, a finding consistent with prior research that has identified education attainment as a significant predictor of retaining mobile devices and mobile applications use [3–5]. This could be attributed to the interest in the use of new technology, the perceived value of using mHealth

applications, and better skills [3]. Other research [3, 4] has identified independent relationships between age and gender and the use of mobile health applications.

The easy use of mobile health applications was found in our study to be a major factor for their use. According to other studies, satisfaction and the perceived ease of use were also to be linked to the desire to use mobile health applications on smartphones [6–9].

Privacy concerns and the misuse of personal data in our study were key roadblocks to using mobile health applications. This finding supports the earlier investigations [10–17]. People are not constantly trusting storing private information online, according to a previous study [18]. Trust is a critical component of any community's success or failure, especially in a virtual world [19].

These concerns are understandable, given that patients' health records on many occasions have been unprotected from hacking [20], and many mHealth applications lack the requisite security protections to safeguard users' health information [21, 22].

In our study, participants' feeling of discomfort and dissatisfaction was found to be independent barrier correlated with the use of mHealth applications. This finding is in line with the Nielsen usability model [23], which rated satisfaction as the top of the five attributes for mobile devices and applications users. The terms "satisfaction" and "comfort" refer to the user's likeability, comfort, and enjoyment [23].

This study has some limitations. The capacity to discern the causal relationship between predictor factors is limited by the cross-sectional study design. Also, the generalizability of the findings cannot be extended to travelers from other countries to Saudi Arabia because the study participants were only from the local community. However, our study offers intuitions into the prevalence and influential issues concerned with the use of mobile health applications that could be useful when planning for effective strategies to promote their use.

5 Conclusion

While the vast majority of the participants were smartphone owners and have Internet access, only a fair percent used mHealth applications. Higher education, easy access, and ease of use were promoting factors for mHealth usage while having concerns regards violation of privacy and discomfort/dissatisfaction on use were barriers for mHealth usage.

6 Recommendations

There is a need to raise public awareness toward the benefits of such mHealth Apps, and develop easy need-based trusted affordable programs with good users support.

References

1. World Health Organization (WHO) (Homepage) (2012) mHealth: new horizons for health through mobile technologies: second global survey on eHealth. ISBN 978 92 4 156425 0. Available at: https://www.who.int/goe/publications/goe_mhealth_web.pdf. Last accessed 2021/8/22
2. United Nations General Assembly (Homepage). Resolution 70/125 (2015) Available at: https://unctad.org/system/files/official-document/ares70d125_en.pdf. Last accessed 2021/9/10
3. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K (2017) Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 19(4):e125
4. Kontos E, Blake KD, Chou WS, Prestin A (2014) Predictors of eHealth usage: insights on the digital divide from the health information national trends survey 2012. J Med Internet Res 16(7):e172
5. Bhuyan SS, Lu N, Chandak A, Kim H, Wyant D, Bhatt J, Kedia S, Chang CF (2016) Use of mobile health applications for health-seeking behavior among US adults. J Med Syst 40(6):153
6. Cho J (2016) The impact of post-adoption beliefs on the continued use of health apps. Int J Med Inform 87:75–83
7. Palos-Sanchez PR, Saura JR, Rios Martin MÁ, Aguayo-Camacho M (2021) Toward a better understanding of the intention to use mHealth apps: exploratory study. JMIR Mhealth Uhealth 9(9):e27021
8. Akdur G, Aydin MN, Akdur G (2020) Adoption of mobile health apps in dietetic practice: case study of diyetkolik. JMIR Mhealth Uhealth 8(10):e16911
9. Peng W, Kanthawala S, Yuan S, Hussain SA (2016) A qualitative study of user perceptions of mobile health apps. BMC Public Health 16(1):1158
10. Liu Y, Yang Y, Liu Y, Tzeng GH (2019) Improving sustainable mobile health care promotion: a novel hybrid MCDM method. Sustainability 11(3):752
11. Boukerche A, Ren Y (2009) A secure mobile healthcare system using trust-based multicast scheme. IEEE J Sel Areas Commun 27(4):387–399
12. Lee KC, Chung N (2009) Understanding factors affecting trust in and satisfaction with mobile banking in Korea: a modified DeLone and McLean's model perspective. Interact Comput 21(5–6):385–392
13. Krebs P, Duncan DT (2015) Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 3(4):e101
14. Atienza AA, Zarcadoolas C, Vaughon W, Hughes P, Patel V, Chou WS et al (2015) Consumer attitudes and perceptions on mHealth privacy and security: findings from a mixed-methods study. J Health Commun 20(6):673–679
15. Kotz D, Gunter CA, Kumar S, Weiner JP (2016) Privacy and security in mobile health: a research agenda. Computer 49(6):22–30
16. Zhou L, Bao J, Watzlaf V, Parmanto B (2019) Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. JMIR Mhealth Uhealth 7(4):e11223
17. König LM, Attig C, Franke T, Renner B (2021) Barriers to and facilitators for using nutrition apps: systematic review and conceptual framework. JMIR Mhealth Uhealth 9(6):e20037
18. Zhang JY, Wu P, Zhu J, Hu H, Bonomi F (2013) Privacy-preserved mobile sensing through hybrid cloud trust framework. In: Cloud computing (CLOUD), IEEE sixth international conference, pp 952–953. Santa Clara, CA, USA
19. El Morr C, Kawash J (2007) Mobile virtual communities' research: a synthesis of current trends and a look at future perspectives. Int J Web Based Commun 3(4):386–403
20. US Department of Health & Human Services (Homepage). Office for Civil Rights (2017) Breach portal: notice to the secretary of HHS breach of unsecured protected health information. Available at: https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf. Last accessed 2021/9/18

21. Papageorgiou A, Strigkos M, Politou E, Alepis E, Solanas A, Patsakis C (2018) Security and privacy analysis of mobile health applications: the alarming state of practice. IEEE Access 6:9390–9403
22. Müthing J, Jäschke T, Friedrich CM (2017) Client-focused security assessment of mHealth apps and recommended practices to prevent or mitigate transport security issues. JMIR Mhealth Uhealth 5(10):e147
23. Nielsen J, Mack R (1994) Usability inspection methods. Wiley, New York, NY

Literature Review of TAM Model Applicable to e-government in Peru's Agricultural Export Sector

Salas Cesar, Vega Hugo, and Rodriguez Ciro

Abstract The research carries out a bibliographic review of state-of-the-art factors that affect the process of adopting technology and the models for its acceptance, focusing specifically on the Peruvian government's information systems so that administrative procedures are faster and more efficient. Competitiveness in productive activities is important to guarantee quality and lower prices in products and services delivered to the economy. The methodology considers the planning review, the conduct review, and the report review; in the process, the existing evidence of methods and models to identify knowledge gaps is summarized. The objective of reviewing the state-of-the-art of TAM models from 2001–2019 was completed, finding 10 models and 36 factors proposed by researchers on this subject, all of these are related to the perceived usefulness and ease of use associated with the use of the system, as well as some elements that complement the adoption process.

Keywords TAM · Technology acceptance model · A literature review · ICT adoption · Factors · Models · E-government · Agricultural export sector

S. Cesar (✉)
Peruvian University of Applied Sciences (UPC), Lima, Peru
e-mail: pcsicsal@upc.edu.pe

S. Cesar · V. Hugo · R. Ciro
Department of Systems and Information Engineering, National University of San Marcos (UNMSM), Lima, Peru
e-mail: hvegah@unmsm.edu.pe

R. Ciro
e-mail: crodriguezro@unmsm.edu.pe

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_19

1 Introduction

For this research, the authors focus on conducting a literature review of a technological acceptance model, which allows technology to be adopted faster in companies, specifically focusing on information systems provided by the Peruvian government so that administrative procedures are faster and allow government to be more competitive.

There are variables identified as critical to adopting technology, such as: "creates an effective governance structure, assertive communication, social influence, early training development, utility, and perceived ease of use" according to [1]. However, they are not part of a model and them there is not a formal way to approach the process of adopting technology systemically. Our contribution consists in a review state of art related to Technology Acceptance Model—TAM to introduce this research to the community.

The article is divided into six sections, in background, submits the literature reviews about the TAM model and factors which affect technology acceptance. In Sect. 3: Research Methodology, details the procedure for performing systematic literature reviews, explaining in three phases: planning, conducting, and reporting the review. Section 4: Analysis and answers the four questions to obtain the results of the research: state of art related to TAM Models. Section 5: Discussion, presents two tables that summarized the results of the research. Finally, in Sect. 6 presents: Conclusion.

2 Background

2.1 Technology Acceptance Model—TAM

The TAM model, known for being a model that allows mapping human constructs when making a decision for or against accepting to use a new technology. It is a model that has been studied for the last 40 years by researchers in the information systems (IS) area and still has a lot of room to continue developing since we do not have overwhelming levels of acceptance of technology and must accept canning as it comes from the factory.

2.2 Peru's Agricultural Export Sector

This sector is strategic for the interests of national policy, not only because it offers 1.5 million jobs among direct and indirect, but also improves the country's trade balance, given the level of agricultural exports of products called traditional and non-traditional, however it is also the second most dynamic sector after mining and

has been growing at sustained rates of 4.5%, contributing directly to the national gross. According to the Ministry of Foreign Trade and Tourism through the Foreign Trade Single Window (VUCE), "the number of companies linked to foreign trade that carry out agricultural merchandise export tasks as of December 2018 are registered: 4560 companies duly constituted with numbers RUC also registered in the National Superintendency of Customs and Tax Administration (SUNAT)" [1].

3 Research Methodology

This research considers the procedure proposed by Keele [2] for "performing systematic literature reviews where consist in 3 phases":

- Phase 1: Planning the review, the studies questions are raised and the revised protocol is defined.
- Phase 2: Conducting the review, in this phase the research plan is executed and the primary studies are selected, according to the inclusion and execution criteria.
- Phase 3: Reporting the review, for this phase, statistics and analysis of documents are used for selecting and discussing.

There are many reasons for undertaking a systematic literature review, the most common reason is:

- To summarize the existing evidence of the benefits and limitations of methods, models, and methodologies.
- To identify gaps in current research to suggest areas for further investigation.
- To provide background to properly position new researching activities.

3.1 Phase 1: Planning the Review

The authors decided to ask themselves the following questions about TAM Models, methods or artifact for Technology acceptance Models.

- Question 1: What is TAM Model?
- Question 2: Which factors influence in the TAM Models?
- Question 3: What is the TAM Model and which factors influence it?
- Question 4: Who are the most representative authors of TAM Models?

Google Scholar, ACM Digital Library, Science Direct, Web of Science, and Others research engine as: ResearchGate, Mendeley have been considered according to the four strings and sources given in Table 1. A list of the resulting studies which have a title, an abstract, a journal name, authors, a published date, a referenced number, and DOI.

Other parameters used in the research were: language, documents, and publication types, period comprised January 2001 and January 2019.

Table 1 Key words string used in search

Source	Key words string search
ACM digital library	String 1 = (technological adoption) or TI = (adoption model) or TI = (technological absorption) or TI = (export sector of agricultural goods))
Web of science	String 2 = (IT adoption) or TI = (IT adoption model) or TI = (technological absorption index))
Google scholar	String 3 = (adoption of technology) or TI = (agricultural products) or TI = (foreign trade, food safety))
Science direct	String 4 = (technology acceptance model (TAM)) or TI = (TAM 2) or TI = (dissemination of innovations) or TI = (ease of use and perceived utility))
Others	

3.2 Phase 2: Conducting the Review

The search engine was applied according to the plan, applying selection criteria from Table 1 and the searching parameter showing at least 42,560 studies.

The next step was applied the selection and exclusion criteria listed in Table 2, obtaining only 500 studies as a result, being the most outstanding criteria in the exclusion of articles related to the use or application of the TAM model and as inclusion criteria only the type documents: the journal.

The following step: Studies were analyzed the papers resulted and adapted to match the four author's questions to get a list of newly filtered, 38 studies detailed in Table 3.

Its shows 23 journals related to the topics of: technology, management, innovation and communication have been identified that coincide with publications on TAM models.

Table 2 Selection and exclusion criteria

Selection criteria	Exclusion criteria
Articles are submitted factors, models, methodologies about: definition, evolution, development TAM model	Articles are submitted factors, models, methodologies about: use or application TAM model
The answer to the authors questions	Elements that are not with TAM model context
Authors of articles were relationships between them by semanticscholar.org	Authors of isolated topics
Articles from sources with impact factor SJR, publication type: journals	Books, theses, posters, proceedings, conferences, technical reports

Table 3 Studies resulting from the application of search with different criteria

Source	Studies resulted from applying search		
	With keywords string	With selection and exclusion criteria	Which give a response to one of four questions
ACM digital library	33,291	340	15
Web of science	2,040	61	7
Google scholar	2,582	55	5
Science direct	4,622	36	3
Others	25	8	6
Total	42,560	500	38

3.3 Phase 3: Reporting the Review

Finally, 38 papers were selected; they were used to build a state of art from Technology Acceptance Model with an Information System perspective. The authors who are responsible for the evolution of the TAM model, since its creation by Fred Davis, who in the 80's was financed by IBM propose this model and subsequently evolves the model, adopting a binary relationship with Viswanath Venkatesh who evolved the model to TAM2 in the 90's with Richard Bagozzi. The participants of Ramayah and Rogers are complementary, giving a necessary innovation approach the 2000's era [3].

4 Finding and Analysis

4.1 Question 1: What is TAM Model?

Known by its acronym in English TAM, it is a theorical model how permitted users evaluate accepting or not for using any technology. So, it is strictly necessary to analyze the process, models, techniques, tools, proposed by ICTs that allow to identify the strengths of the TAM model and its applicability to the acceptance needs of new technologies to the intended target audience.

The analysis starts assessing the problem through the disciplines of human knowledge as sociology and psychology, responsible from people absorb, process and take a decision, according to authors "in its main currents: sociocultural and psychosocial, seeks to clarify what are the basis to analyze human behavior in the social environment, taking into account how the latter influences the beliefs—each individual has—in relation to the use of technology" [1].

Below are the most representative theories and models shown by its main creators such as: Venkatesh et al. [4]. It should be noted that TAM was applied very frequently in technology adoption processes: e-commerce [5, 6].

Table 4 Evolution of models/theories that affect the technology acceptance model—TAM

Model/theory	Date	Sources
Theory of reasoned action (TRA)	70's	[7]
Theory of planned behavior (TPB)	80's	[8]
Theory of planned behavior (TPB) and decomposed theory of planned behavior (DTPB)	1991 1995	[9] [10]
Technology acceptance model (TAM)	1996–2006	[2, 11]
Technology acceptance model (TAM2)	2000	[12]
Technology acceptance model (TAM3)	2008	[13]
Unified theory of acceptance and use of technology (UTAUT)	2003	[4]
Unified theory of acceptance and use of technology (UTAUT2)	2012	[11]

4.2 Question 2: Which Factors Influence in the TAM Models? and Question 3: What is the TAM Model and Which Factors Influence It?

This answer was wide developed in Background section, when explain in which consist of: Technology Acceptance Model (TAM). TAM model evolution over the time was detailed in Table 4.

According to McCoy et al. [14] as a reference valid to explain the evolution of TAM.

4.3 Question 4: Who Are the Most Representative Authors of TAM Models?

The most representative authors of the TAM show the 25 authors of the technology acceptance model. From now on, Venkatesh and Bagozzi make a break in their research accompaniment working independently evolving the model from their perspectives and a group of independent researchers. Among these important authors are Wynne Chin (China) and Icek Ikzen (Indian) who contribute favorably to the evolution of the TAM model reaching TAM3 to finally remain in the Unified Theory of the Acceptance of the Use of Technology (UTAUT) which now, has some dominance in the world orb.

The most outstanding thing about this graph is that: it not only shows the evolution in time of the TAM model,—but also the authors' geopolitical pressures from their nations of origin: Davis and Bagozzi (USA), Venkatesh (Indian), Wynne Chin (China) who strive for leadership in the development of these models opt to become the leaders in the adoption technology.

4.4 To Sum up

The Technological Acceptance Models (TAM), a product of the bibliographic research was analyzed in two types of research: the initial ones that are of the basic type responsible for developing and evolving the model and those which use the models by applying different emerging technologies applicable to the geographic sectors and regions in the world.

In Table 5, the most referenced model is detailed from a state-of-the-art of TAM Models- ordered at the time of evolution—they were launched into the world of science.

Below, The TAM models are listed in Table 6, which contain moderators: predictors or constructs or factors that are decisive in obtaining results when applying the

Table 5 Technology acceptance models without moderators—state of art (1/2)

ID	Constructs/factors	Model/theory	Source
RA0	Relative advantage	Innovation diffusion theory (IDT)	[15]
COT	Compatibility		
COM	Complexity		
TRI	Trialability		
A00	Attitudes	Theory action reasoned (TAR)	[8]
SN0	Subjective norms		
A0	Attitudes	Theory of planned behavior (TPB)	[16, 17]
SN0	Subjective norms		
PBC	Perceived behavioral control		
IM0	Intrinsic motivation	The motivational model (MM)	[16, 18]
EU0	Ease of use	The technology acceptance model (TAM)	[18]
PU0	Perceived utility		
JF0	Job—fit	The PC utilization model (MPCU)	[16, 17]
C00	Complexity		
LTC	Long-term consequences		
ATU	Affect toward use		
SF0	Social factors		
FC0	Facilitating conditions		
SE0	Self-efficacy	The cognitive social theory (SCT)	[19]
SR0	Self-regulation		
OL0	Observational learning		
OE0	Outcome expectations		
BC0	Behavioral capability		
REF	Reinforcement		

Table 6 Technology acceptance models, with moderatos—state of art (2/2)

ID	Moderators: factor/predictor/construct	Model/theory	Source
SN0	subjective norm	TAM2, as evolution of TAM	[12, 18]
VOU	Voluntariness of use		
I00	Image		
PE0	Perceived enjoyment	TAM3	[13]
PEC	Perception of external control		
OU0	Objective usability		
CA0	Computer anxiety		
CP0	Computer playfulness		
EE0	Effort expectancy	Unified theory of acceptance of technology (UTAUT)	[4]
PEX	Performance expectancy		
SI0	Social influence		
FC0	Facilitating conditions		

TAM model, being the third important variable in obtaining results, this variable is known as: effect modifier.

5 Discussion

This research provides insights into the TAM model, although no shows studies specifically in agro-export sector, there were research of using this model in e-government. Considering the existence of studies that analyze the TAM models and identify important factors for the proper development of a framework to subsequently evaluate the performance of the model in the face of an e-government project where all the stakeholders in its application were considered [20]. The studies used are from research area published in the past eighteen years, 2001 to 2019 where collect the most popular models envelope TAM model. It is seen from "the most commonly used research models for studies in this topic using methodology proposed by Keele, has been the quantitative research" [20].

According to Singh [20] "Recently, new factors such as equality, risk and economic value, have become important and part of the assessment tools of TAM models". So Peruvian e-government service that supports administrative procedures for import and export companies from the agricultural merchandise turn around faster and more efficient when use a model with factors as: trust and risk perceived are more valued by stakeholders from this economy sector [21].

6 Conclusion

Nowadays, more than ever, several steps are needed to adopt technology in the most efficient and formal way. The way to do it, is by researching and taking knowledge of the existing models, methods, procedures; as well as the factors, constructs or elements that influence them in order to evaluate the model allowing the users to adopt more efficient technology to the reality of the needs of their: Organization, economic sector or ecosystem. This investigation complies with the objective of making a review of state-of-the-art through TAM model evolution, which factors make them up and which authors are the most predominant in the investigation of this topic, finding in this article, a summary of the models and factors that have been proposed by researchers on the subject when most of them categorized as factors coming from TAM Models with moderators.

According to the bibliographic references found, there are evidence that criticize the TAM model, which need to be extended [22]. It is showing a gap among the models found and the adoption needs that organizations submit, finding space to carry out future research where complementary models to the TAM were formulated that address factors. According to McKnight, V., found a trust as a factor needs to be studied [23].

References

1. Arbaiza CES (2018) Critical variables for success in the technology adoption process in the framework of digital transformation. In: ACM international conference proceeding series. https://doi.org/10.1145/3277139.3277163
2. Keele S (2007) Guidelines for performing systematic literature reviews in software engineering. In: Technical report, Ver. 2.3 EBSE technical report. EBSE
3. Pantano E, Di Pietro L (2012) Understanding consumer's acceptance of technology-based innovations in retailing. J Technol Manag Innov. https://doi.org/10.4067/S0718-27242012000400001
4. Venkatesh, Morris, Davis, Davis (2003) User acceptance of information technology: toward a unified view. MIS Quarterly. https://doi.org/10.2307/30036540
5. Chandio FH, Burfat FM, Abro AA, Naqvi HF (2017) Citizens' acceptance and usage of electronic-government services: a conceptual model of trust and technological factors. Sindh Univer Res J Sci Series. https://doi.org/10.26692/surj/2017.09.36
6. Ha S, Stoel L (2009) Consumer e-shopping acceptance: antecedents in a technology acceptance model. J Bus Res. https://doi.org/10.1016/j.jbusres.2008.06.016
7. Fishbein M, Ajzen I (1975) Chapter 1. belief, attitude, intention, and behavior: an introduction to theory and research. Reading, MA, Addison-Wesley
8. Ajzen I, Fishbein M (1980) Understanding attitudes and predicting social behaviour. Eagle Wood Cliffs, NJ, Prentice- Hall. Englewood Cliffs
9. Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process. https://doi.org/10.1016/0749-5978(91)90020-T
10. Todd PA, Taylor S (School of Business Queen's University Kingston, O., & Canada K7L3N6) (1995) Understanding information technology usage: a test of competing models. Inf Syst Res
11. Venkatesh V, Thong JYL, Xu X (2012) Venkatesh_Thong_Xu_MISQ_forthcoming (Gender Age Experience). MIS Quarterly

12. Venkatesh V, Davis FD (2000) Theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci. https://doi.org/10.1287/mnsc.46.2.186.11926
13. Venkatesh V, Bala H (2008) Technology acceptance model 3 and a research agenda on interventions. Decis Sci. https://doi.org/10.1111/j.1540-5915.2008.00192.x
14. McCoy S, Galletta DF, King WR (2007) Applying TAM across cultures: the need for caution. In Euro J Inf Syst. https://doi.org/10.1057/palgrave.ejis.3000659
15. Rogers EM (1983) Difussion of innovations. In: The Free Press
16. Dulloo R, Mokashi J, Puri MM (2014) Exploring technology acceptance theories and models—a comparative analysis. MIT-SOM PGRC KJIMRP 1st international conference
17. Thompson RL, Higgins CA, Howell JM (1991) Personal computing: toward a conceptual model of utilization of personal computers personal computing: toward a conceptual model of utilization1. Source: MIS Quarterly
18. Davis FD, Bagozzi RP, Warshaw PR (1989) User acceptance of computer technology: a comparison of two theoretical models. Manage Sci. https://doi.org/10.1287/mnsc.35.8.982
19. Bandura A, Cervone D (1986) Differential engagement of self-reactive influences in cognitive motivation. Organ Behav Hum Decis Process. https://doi.org/10.1016/0749-5978(86)90028-2
20. Singh H, Grover P, Kar AK, Ilavarasan PV (2020) Review of performance assessment frameworks of e-government projects. Trans Gov: People, Proc Policy 14(1):31–64. https://doi.org/10.1108/TG-02-2019-0011
21. Arbaiza CS, Huerta HV, Rodriguez CR (2021) Contributions to the technological adoption model for the peruvian agro-export sector. Int J E-Adopt (IJEA) 13(1):1–17. https://doi.org/10.4018/IJEA.2021010101
22. Pan YC, Jacobs A, Tan C, Askool S (2018) Extending technology acceptance model for proximity mobile payment via organisational semiotics. IFIP Adv Inf Commun Technol. https://doi.org/10.1007/978-3-319-94541-5_5
23. McKnight D, Choudhury V, Kacmar C (2002) The impact of initial consumer trust on intentions to transact with a web site: a trust building model. J Strat Inf Syst J Strat Inf Syst. https://doi.org/10.1002/bip.360350308

Internet of Things (IoT) Adoption: Challenges and Barriers

Abdulrahman S. Alenizi and Khamis A. Al-Karawi

Abstract Today, the Internet is one of the most expanding and changing technologies, and it has grown popular all around the globe. The Internet of Things (IoT) is a system that includes a device, a sensor, a network, cloud storage and an application. Every interface to communicate with another device over the Internet to share information and achieve specific objectives Internet of Things (IoT) is a new future technology that is gaining traction in various fields around the world. Kuwait is one of the nations in the planning stages of expanding IoT development, comparable to other countries with rising IoT application development. However, owing to several obstacles and challenges in integrating IoT devices, it was not simple to design IoT devices. This article highlights IoT's key concerns, barriers and solutions to these issues. IoT's future trends and uses were also briefly explored in this article to acquire a more in-depth understanding of IoT equipment.

Keywords Internet of things · IoT adoption · Implementation · Challenges and barriers · Kuwait

1 Introduction

The term "Internet of Things" refers to network devices' ability to detect and collect data and then send it over the Internet to be analysed and used for various fascinating purposes. Furthermore, IoT encompasses protocols, software and communication technologies that make these objects possible to communicate, interact and exchange data and commands without the need for human interaction [1]. To present, several intelligent instruments on the market are directly connected to the Internet. The number of gadgets continues to grow in tandem with technological advancements. Innovative technologies are a component of nature that makes up the Internet of

A. S. Alenizi (✉)
Business School, University of Salford, Manchester, UK
e-mail: a.alenizi@edu.salford.ac.uk

K. A. Al-Karawi
University of Diyala, Diyala, Baqubah, Iraq

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_20

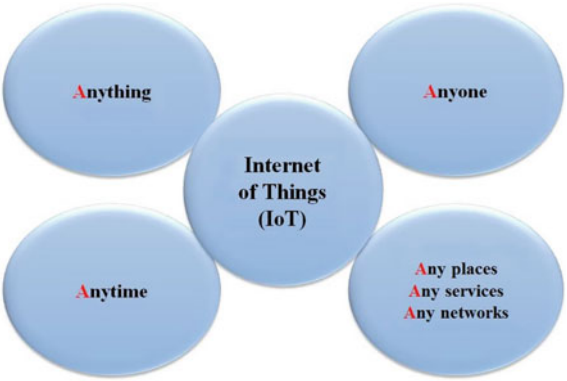


Fig. 1 Internet of Things concept

Things which is undergoing a digital revolution [2]. IoT refers to a connected device that permits connection with other devices, after which data is taken and processed into valuable and relevant information [3, 4]. RFID, remote wireless communication, real-time localisation and sensor networks are typical IoT technologies. Smartphones, tablets and laptop computers are examples of smart gadgets and technology considered Internet of Things components [5]. With the Internet of Things (IoT) technologies, many devices connected can be used to assist the report in recording the identifier and location via wired or wireless connections [6–8]. The Internet of Things is depicted in Fig. 1 as a collection of descriptions [8]. The Internet of Things is a brand-new method of connecting to the Internet. Because they can provide information about themselves to others, objects can identify themselves and gain intelligent behaviour via the Internet of Things by making or enabling relevant choices [9]. These objects may access data gathered by other items or be contributed to other services [9]. Figure 1 shows how the IoT will allow anything to interact with the Internet and from any location to provide any assistance to anyone over any network. This concept will lead to the creation of new applications, such as intelligent vehicles and smart homes, that will provide various services, including alerts, security, energy savings, automation, communication, computing and entertainment. Intelligent environments will be much closer to being implemented by 2020 if IoT technology is developed, tested and deployed [10]. A highly decentralised wide range of resources connected by a dynamic network of networks will soon include people, machines, smart objects, the environment and platforms. This paper aims to provide an overview of IoT applications in Kuwait, as well as future technologies and issues. On the rest of the paper: in the Sect. 2, we will look at IoT applications. In Sect. 3, the Internet of Things challenges and barriers; In Sect. 4, Internet of Things architecture; In Sect. 5, results and discussion. Finally, the conclusion.

2 IoT Applications

The IoT has numerous applications in human life, including making life easier, safer and more intelligent. Smart cities, smart homes, transportation, energy and intellectual environments are just some applications. Integration of IoT with a cognitive environment is required to expand the capabilities of intelligent objects and allows the user to monitor the environment remotely (Fig. 2).

2.1 Smart Cities

Innovative developments have benefited many major cities, including Seoul, New York, Tokyo, Shanghai, Singapore, Amsterdam and Dubai. With the current rate of innovation in constructing smart cities, including IoT technology in city development, smart cities may still be considered future cities and intelligent living. Smart cities need meticulous planning and cooperation from governments and residents to incorporate Internet of things technology in all sectors. Cities can benefit from the Internet of Things on various levels, as shown in Fig. 3, including improving infrastructure, expanding public transportation, reducing traffic congestion and keeping residents safe, healthier and more involved in their communities [11].

The Internet of Things will make cities smarter by connecting all city systems such as transportation, healthcare, weather monitoring and so on and allowing people in any location to access databases of airports, railways and transportation tracking that operate under specific protocols via the Internet.

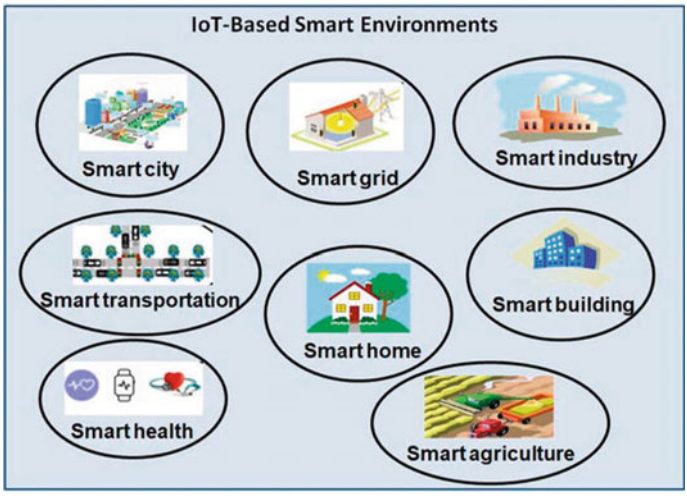


Fig. 2 IoT-based smart environments

Fig. 3 Smart cities aspects

2.2 Smart Buildings and Homes

Wi-Fi technologies have been used in home automation systems. Televisions, cell phones and other gadgets are commonly connected via Wi-Fi. It is possible to control various devices in a home using the Internet of Things concept. A portable "controller" for network-connected devices is also provided by mobile devices. They can both serve as gateways for the Internet of Things [7, 12–15]. Intelligent energy management in buildings can be achieved through the use of wireless sensor networks (WSNs) with Internet of Things integration. Facilities managers will be able to control energy consumption and procurement, as well as perform maintenance and repairs on building systems. Figure 4 depicts some of the most intriguing Internet of Things (IoT) applications in smart buildings and homes.

2.3 The Smart Grid and Smart Energy

A sensor network is a system that connects access and control to provide intelligent energy management [16]. A sensor network that incorporates modern information and communication technologies (ICTs) into the electricity network will allow real-time, two-way communication between suppliers and consumers, resulting in a more dynamic interaction on energy flow and more sustainable and efficient transmission and distribution of electricity. The smart grid is one of the most important applications of the Internet of Things. Smart grids can manage a wide range of applications, including manufacturing, solar power, nuclear power, automobiles, hospitals and city power management. There are numerous benefits for customers, including lower

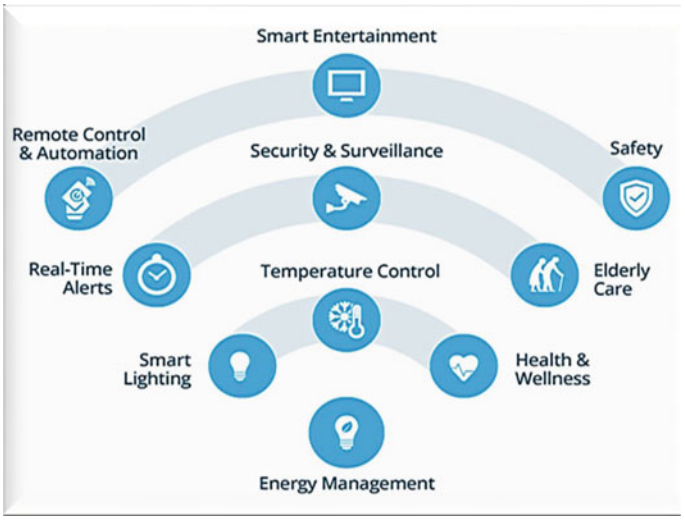


Fig. 4 Smart home and building applications

utility bills, more efficient use of electricity at home and the integration of renewable energy and green technologies [17, 18] (Fig. 5).

2.4 Smart Health

IoT monitoring technology allows for the continuous monitoring of hospitalised patients whose physiological conditions must be maintained. Smart health sensors are utilised to capture extensive physiological data, which is then evaluated and stored by gateways and the cloud before being wirelessly sent to caregivers for additional analysis and evaluation, as illustrated in Fig. 6. It eliminates the need for a health expert to visit the patient regularly to check their vital signs instead of delivering a continuous automatic flow of information. This approach simultaneously enhances care quality via continual attention and decreases care costs by lowering the cost of conventional care methods and data collecting and analysis [7].

Several more people worldwide are in poor health due to a lack of easy access to proper health monitoring, and they may be classified as critical situation patients. Compact, reliable wireless solutions linked via the Internet of Things, on the other hand, allow monitoring to reach these patients. These technologies could be used to record patient health data from various sensors securely, assess the data using complex algorithms and then share it with medical specialists via wireless links so they can make the appropriate clinical guidelines [19].

Fig. 5 Smart grid applications

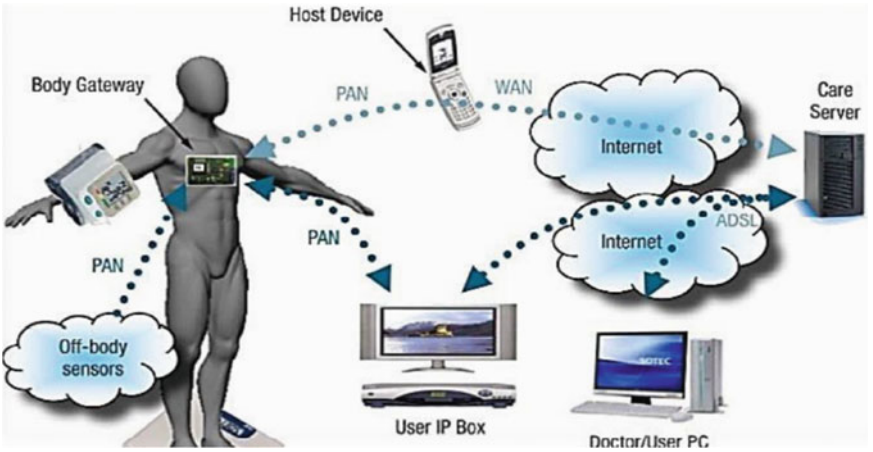


Fig. 6 Smart healthcare concept

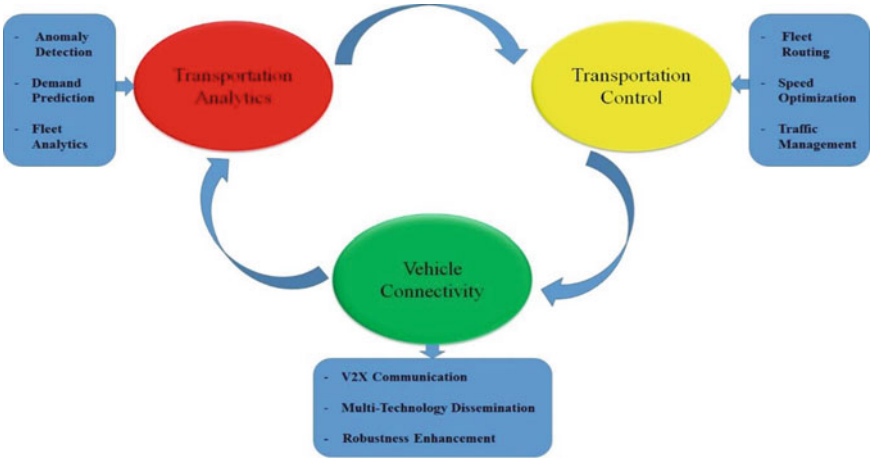


Fig. 7 Smart transportation aspects

2.5 Intelligent Public Transport and Mobility

Road condition monitoring and alarms are an important IoT application. Multi-technology dissemination is used to manage vehicle routing, speed control and traffic management. Transport can also benefit from IoT, such as electric cars, which are gaining popularity as a way to reduce fuel costs and the impact of global warming. Lithium-ion batteries have been the subject of extensive research funded by a number of governments [20] (Fig. 7).

2.6 Smart Manufacturing and Smart Factory

Artificial intelligence, machine learning, expertise labour mechanisation and M2M communication are incorporated into the manufacturing process [21]. Allowing low-emission and incidental manufacturing can improve safety and protect the environment. Manufacturing is becoming "smarter" as a result of advancements in the interaction of machines and other items and the resulting shift in decision-making from people to technological systems [13]. Physical cryptographic systems linked to the Internet are at the heart of Industrial 4.0, also referred to as an intelligent industry. The "industrial Internet of Things" will enable "smart factory" and "smart manufacturing" to be fully realised through M2M communications enabled by big data. The country's research and innovation policy will be refocused on forward-thinking initiatives related to scientific and technological breakthroughs following the launch of the high-tech strategy 2020 initiative. This is depicted in Fig. 8.

Fig. 8 Smart factory (industry 4)

2.7 Smart Environment

Human existence is profoundly influenced by the natural environment. An unhealthy environment can harm all of these, including people, animals, birds, fish and plants. Environmental pollution and waste resources have been the subject of numerous studies [7]. Industrial and transportation wastes and irresponsible human actions are everyday causes that affect the ecosystem, making it challenging to create a healthy environment. Monitoring the environment is essential to assess the current state and make correct life decisions based on data collected from monitoring systems. Management must have efficient resource consumption and use and reduce factory and vehicle wastes. Monitoring and waste management give vast data to push governments or healthy environment groups to enforce health standards [22]. An intelligent environment is an essential part of our daily lives, as shown in Fig. 9, since it offers numerous services and remedies for a wide range of environmental uses, including water and air pollution, weather and radiation monitoring, waste management, natural disasters and a variety of other ecological indicators. Smart environmental devices that integrate with Internet of Things (IoT) technology are currently being developed. There are several Internets of things applications for the environment, divided into environmental resource management and environmental quality and protection

Fig. 9 Architecture of the Internet of Things [27]

management. Resource management has an impact on all natural resources, including animals, plants and forests, birds and fish, coal, petroleum, land, freshwater, air and heavy metals such as gold, copper and iron. IoT can facilitate communication between these resource sensors and research and monitoring centres, allowing for more informed resource allocation decisions [23].

2.8 Academic

Kuwait is preparing to introduce a new having to learn procedure known as massive open online courses, which promotes open learning (MOOC). This way of education enables anyone in the society to connect the information and learning reference anywhere at time and from any place, while also sharing knowledge.

3 Internet of Things Challenges

Even though the stages of the business described above are fascinating in providing technology for bright everything, there are some costs associated with putting the Internet of Things concept into practice. In addition to the assumption that the technology must be low-cost and applicable to a wide range of things, IoT faces a slew of other issues [24], including:

- **Scalability**: To achieve scalability and efficiency, new processes and methodologies must be developed.
- **Self-Organising**: Smart objects should not be treated the same way as computers. Mobile items must connect, arrange and configure themselves to their surroundings.
- **Data volumes**: Sensor networks, logistics and large-scale networks are examples of Internet of Things application scenarios. This phenomenon necessitates new operational mechanisms and technology for data storage, processing and administration.
- **Data interpretation**: To help savvy item users, it is important to understand the local situation as accurately as possible. Services must derive generalisable inferences from analysed sensor data to benefit from the heterogeneous data.
- **Interoperability**: The Internet of Things' information, processing and communication capabilities vary by smart item. Constraints on smart items include energy availability and communication bandwidth requirements. Common standards are required for these objects to communicate and collaborate.
- **Automatic Discover**y: In dynamic contexts, relevant services for items must be recognised automatically, which necessitates good semantic ways of defining their functioning.

- **Software complexity**: More comprehensive software architectures on the network and background servers will be required to manage intelligent objects and deliver services.
- **Security and privacy**: In the IoT, specific services must be accessed or prevented from connecting with other things, and commercial transactions involving smart items must be safeguarded from prying eyes.
- **Power supply**: RFID transceivers do not need power, but their range and effectiveness are limited. Future embedded systems processing and communication modules should be smaller. Energy conservation is a consideration in both hardware and software design.
- **Wireless communications**: Newer wireless technologies like ZigBee use less power but have less bandwidth. GSM, UMTS, Wi-Fi and Bluetooth must be replaced by faster and more efficient wireless standards.

4 Architectures of IOT

There are already many devices that connect to the Internet, such as robots, drones and sensors. These gadgets are primarily part of an ecosystem known as the Internet of Things (IoT), which epitomises today's digital revolution. This ecosystem comprises devices, sensors, networks, cloud storage and apps that work together to assist enterprises in improving their strategic posture [25, 26]. Figure 9 depicts the Internet of Things architecture.

4.1 Sensing Layer (Hardware)

The sensing layer includes wireless sensor devices, embedded systems, RFID readers and other sensors in various formats. In the sensor field, hardware that supports information identification and storage, software that collects data from the sensor network, communication and data processing control are all critical components.

4.2 Networking and Communication

First data management layer in IoT is Internet backbone. It includes data routing, publishing, subscribing and cross-platform execution. It must be able to change the network in order to automatically identify and map things. Prioritise power consumption, network management, energy efficiency and service quality (QoS).

4.3 Service Layer

The supplier layer is in charge of maintaining device and data tracking. This layer is responsible for data filtering, information discovery, integration services, service components and access to IoT technologies.

4.4 Interface Layer

As RFID technology advances, many new applications will be included in the "Internet of Things" umbrella. Each IoT device brings benefits to our daily lives, such as the ability to assist individuals, businesses and communities. People may come across this new concept in various settings, including health, safety, business and everyday activities.

5 Results and Discussion

Most countries worldwide develop intelligent city programmers focused on energy, water, transportation, buildings and governance. Five countries lead 84 per cent of the efforts: the USA, Europe, Japan, China and Korea. Other nations, such as India, Australia, South Africa, Canada and Singapore, are also developing digital and linked infrastructure. Millions of items will be wiser and more intelligent by the end of this decade, dramatically improving people's lives. Several organisations are working to standardise IoT-related technology to provide more effective and safe implementations. The creation of IoT applications based on client needs changes in response to user expectations.

6 Conclusion

Kuwait is not alone among Southeast Asian countries with the necessary infrastructure for IoT adoption. Kuwait's government intends to make IoT available across the country by 2015. Kuwait also has policies and objectives for developing IoT devices linked to other countries' development. The IoT has emerged as a predicted trend in the information industry development, resulting in a higher standard of living. Big data has become more complex in detection, communication, control and awareness production, but its growth will continue. Improving these challenges in the coming years will be a powerful and daring step for commercial, industrial and academic networking and communication. Adding a facility to handle unified, seamless, universal Internet connectivity, standardisation and interoperability are all

possible enhancements. Research is also needed in energy sustainability, privacy and security.

Acknowledgements The authors would like to express their gratitude to Kuwait's Public Authority for Applied Education and Training for their assistance with this project. The editors and anonymous reviewers deserve thanks for their insightful comments and suggestions, which helped to improve the paper.

References

1. Zeng D, Guo S, Cheng Z (2011) The web of things: a survey. J Commun 6:424–438
2. Alameri I (2018) MANETS and Internet of Things: the development of a data routing algorithm. Eng Technol Appl Sci Res 8:2604–2608
3. Mineraud J, Mazhelis O, Su X, Tarkoma S (2016) A gap analysis of Internet-of-Things platforms. Comput Commun 89:5–16
4. Singh D, Tripathi G, Jara AJ (2014) A survey of Internet-of-Things: future vision, architecture, challenges and services. In: 2014 IEEE world forum on Internet of Things (WF-IoT), pp 287–292
5. Thota C, Sundarasekar R, Manogaran G, Varatharajan R, Priyan M (2018) Centralized fog computing security platform for IoT and cloud in healthcare system. In: Fog computing: breakthroughs in research and practice, ed: IGI global, pp 365–378
6. Kamilaris A, Pitsillides A (2016) Mobile phone computing and the internet of things: a survey. IEEE Internet Things J 3:885–898
7. Alenizi AS, Al-karawi KA (2022) Cloud computing adoption-based digital open government services: challenges and barriers. In: Proceedings of sixth international congress on information and communication technology, pp 149–160
8. Al-Karawi K (2018) Robust speaker recognition in reverberant condition-toward greater biometric security. University of Salford
9. Ezechina M, Okwara K, Ugboaja C (2015) The Internet of Things (Iot): a scalable approach to connecting everything. Int J Eng Sci 4:09–12
10. Misra G, Kumar V, Agarwal A, Agarwal K (2016) Internet of things (IoT)–a technological analysis and survey on vision, concepts, challenges, innovation directions, technologies, and applications (an upcoming or future generation computer communication system technology). Am J Electric Electron Eng 4:23–32
11. Murray A, Minevich M, Abdoullaev A (2011) Being smart about smart cities. Searcher 19:20
12. Fatima F, Husain BN, Azharuddin M, Mabood MA et al (2015) Internet of Things: a survey on architecture, applications, security, enabling technologies, advantages and disadvantages. Int J Adv Res Comput Commun Eng 4:498–504
13. Al-Karawi KA, Mohammed DY (2021) Improving short utterance speaker verification by combining MFCC and entrocy in noisy conditions. Multimedia Tools Appl 80:22231–22249
14. Al-Karawi KA, Al-Noori AH, Li FF, Ritchings T (2015) Automatic speaker recognition system in adverse conditions-implication of noise and reverberation on system performance. Int J Inf Electron Eng 5:423
15. Al-Karawi KA, Mohammed DY (2019) Early reflection detection using autocorrelation to improve robustness of speaker verification in reverberant conditions. Int J Speech Technol 22:1077–1084
16. Manivannan T, Radhakrishnan D, Preventive model on quality of service in iot applications
17. Miceli R (2013) Energy management and smart grids. Energies 6:2262–2290
18. Al-Karawi KA (2020) Mitigate the reverberation effect on the speaker verification performance using different methods. Int J Speech Technol 1–11

19. Zeinab KAM, Elmustafa SAA (2017) Internet of things applications, challenges and related future technologies. World Sci News 2:126–148
20. Al-Karawi KA, Li F (2017)Robust speaker verification in reverberant conditions using estimated acoustic parameters—a maximum likelihood estimation and training on the fly approach. In: 2017 seventh international conference on innovative computing technology (INTECH), pp 52–57
21. Benioff M (2015) Industrial internet of things: unleashing the potential of connected products and services, January edn
22. Jiang P, Xia H, He Z, Wang Z (2009) Design of a water environment monitoring system based on wireless sensor networks. Sensors 9:6411–6434
23. Al-Karawi KA, Ahmed ST (2021) Model selection toward robustness speaker verification in reverberant conditions. Multimedia Tools Appl, pp 1–18
24. Mishra S, Tripathy HK, Mishra BK, Mohapatra SK (2018) A succinct analysis of applications and services provided by IoT. In: Big data management and the internet of things for improved health systems, IGI Global edn, pp 142–162
25. Medaglia CM, Serbanati A (2010) An overview of privacy and security issues in the internet of things. Internet of Things, pp 389–395
26. Mohammed DY, Al-Karawi KA, Husien IM, Ghulam MA (2019) Mitigate the reverberant effects on speaker recognition via multi-training. In: International conference on applied computing to support industry: innovation and technology, pp 95–109
27. Al-Fuqaha A, Guizani M, Mohammadi M, Aledhari M, Ayyash M (2015) Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun Survy Tutorials 17:2347–2376

Adoption of Information and Communication Technologies in the Agricultural Sector

Sussy Bayona-Oré and Rafael Villon

Abstract Today, information and communications technologies (ICTs) are present in all sectors, and the agricultural sector is no exception. Access to information is important for gaining knowledge and making better decisions. Several factors determine the adoption of ICTs in the agricultural sector. This article presents the literature review on the factors that condition the adoption of ICTs in the agricultural sector. A total of 17 scientific articles were reviewed and analyzed. The results of the review show that some factors mentioned in the literature are education, ICT availability, age of farmers, subjective norm, and farm size. The models that explain the technology acceptance used in the studies were technological acceptance model (TAM) and unified theory of acceptance and use of technology (UTAUT) and theory that explains human behavior such as theory of planned behavior (TPB). In future adoption model proposals, the factors, models, and theory can be taken into account.

Keywords Adoption · Agriculture · ICT · TAM · TPB

1 Introduction

Today, organizations are investing in the acquisition of information and communications technologies (ICTs) in order to improve their processes and provide better services and products to their customers. Agriculture could also take advantage of this widespread use of ICTs, which could represent an opportunity for growth in this sector, improving its inherent processes. Also, the agriculture sector plays an important role in the economy of a country [1].

Emerging technologies such as machine learning have been incorporated into the sector to support farmers' decision-making in the prediction of agricultural product

S. Bayona-Oré (✉)
Dirección de Investigación, Universidad Autónoma del Perú, Lima, Peru
e-mail: sbayonao@hotmail.com

R. Villon
Universidad San Martin de Porres, Lima, Peru

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_21

prices [1, 2]. This demonstrates the importance and benefits that the application and use of ICT can provide a significant contribution to the agricultural sector.

Likewise, the incorporation of ICTs in farmers' livelihood activities could increase their productivity, access to new markets with competitive prices, access to climatological and meteorological information, etc. Despite the benefits offered by ICTs, ICT adoption is still limited by poor infrastructure and access to connectivity [3].

Several studies have been conducted with the purpose of proposing ICT adoption models in the agricultural sector. Among the models that explain the technology acceptance, we can mention the technological acceptance model (TAM) [4], unified theory of acceptance, and use of technology (UTAUT) [5] or theories that explain the human behaviors such as theory of planned behavior (TPB) [6]. Perceived usefulness and attitude toward behavior are two constructs of TAM and predictors of behavioral intention. The constructs of TPB are attitude toward behavior, subjective norms, and perceived behavior control. Some constructs of UTAUT are performance expectancy, effort expectancy, social influence, and facilitating conditions. In the adoption of ICTs, aspects such as access and connectivity [7, 8] are mentioned.

The aim of this article is conducting a literature review in order to know the factors that should be taken into account when proposing a model for ICT adoption in the agricultural sector. This article is a contribution to knowledge and serves as a basis for future research.

The article is structured in five sections including the introduction. Section 2 presents the background. Section 3 presents the methodology. Section 4 presents the results. Finally, Sect. 5 presents the conclusions.

2 Technology Adoption Models in the Agricultural Sector

Growth in the agricultural sector helps diversify the economy and reduce dependence on non-renewable extractive industries and can be a major driver of poverty reduction [9]. Therefore, this sector is very valuable for the growth of the economy, and if the sector grows, then it will generate more jobs. Farmers are the basic unit involved in developing and sustaining the agricultural sector. In this sense, it is important to review the models of technology adoption and identify which constructs can influence the adoption of ICTs by farmers, as well as their use and behavior within the associations. Some of the most important theoretical frameworks are the theory of planned behavior (TPB), the technology acceptance model (TAM), and the unified theory of acceptance and use of technology (UTAUT) [10] among others.

2.1 Adoption of Technologies in the Agricultural Sector

Figure 1 shows the concurrence network map of the keywords of the publications identified in the SCOPUS database using the search chain:

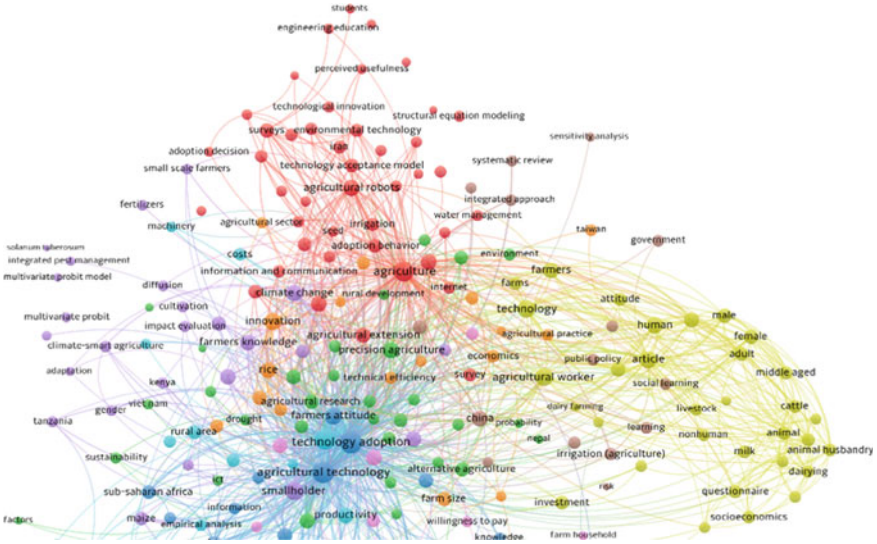


Fig. 1 Map on ICT adoption in agriculture using [11]

(TITLE (ICT OR technology OR tic OR "information communication technology") AND TITLE (adoption OR acceptance) AND TITLE (agriculture OR farm OR "agricultural worker" OR agricultural)).

Figure 1 shows that there are publications with topics associated with the adoption of information technologies in the agricultural sector. The topics that are closely related are technology in agriculture, adoption of technologies, farmers' attitudes, precision in agriculture, among others. In cluster 1, some topics are related to robots in agriculture, technologies in agriculture, Internet, innovation in technology, technology in agriculture, food production, and decision-making.

The topic of technology acceptance and research on people's behavior are aspects that are related to people's attitude toward technology adoption. Technology acceptance model (TAM) and theory of planned behavior (TPB) appear in cluster 1.

With regard to collaboration between countries (Fig. 2), it is observed that the United States, China, Australia and the United Kingdom have the largest number of studies. It is observed that despite the geographical distance between Australia and the United Kingdom and the United States, they are intellectually close. Similar situation between Australia and India, they are intellectually close.

3 Methodology

To begin with, the need for the research and the research question to be answered were established. The research question is

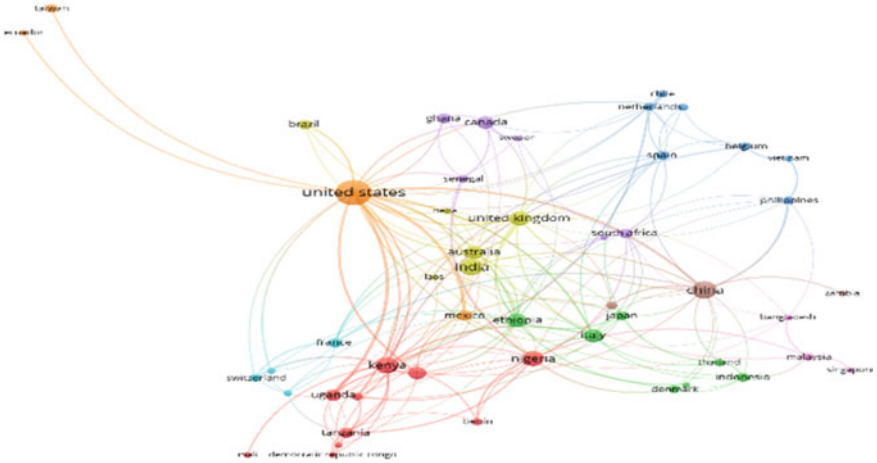


Fig. 2 Countries and their interest in the adoption of technologies in the agricultural sector [11]

What are the factors that contribute to the adoption of ICTs in the agricultural sector?

Next, the key words for the search and the databases to be used to search for scientific articles were defined. Inclusion and exclusion criteria were also determined. The scientific articles considered were published in the period 2010–2022.

The process to select the articles was to review the title, the abstract, and finally, the complete article. Excel sheets were used for data extraction. In the first stage, relevant information was extracted from the articles such as article data, country model used, constructs, proposed model, results, among others. In a second stage, the similar constructs were grouped together.

4 Results

A total of 17 articles were selected. The results of the review show that the authors use different denominations for a construct. From this new universe of constructs, the frequency of repetition per author was determined and taking as a reference, a repetition value equal to or greater than 2. A total of 17 factors were taken into account in order of priority, which were considered to be the most relevant for the research.

The results of the literature show that the applications of acceptance models have been applied to the agricultural sector for the adoption of technologies in areas such as e-commerce, smart cell phones, e-mail, information systems, among others.

Among the main variables found for the study are (1) farmers' education, (2) availability and access to ICT, (3) age of farmers, (4) subjective norm, (5) farm size, (6) perceived usefulness, and (5) attitude toward the use of ICT among others. Table

1 presents the constructs; the first column presents the terms used in the article, and the column 2 presents the author.

The education construct is important in the process of adopting ICTs [7, 7, 12–17], especially digital education, which allows farmers to have greater capacities for the effective management of ICTs [12]. Pivote et al. [7] found that farmers' lack of education, skills, and abilities are barriers to the use of ICT in agriculture. Training activities for farmers [3] can help achieve these skills [12]. Education has a positive influence on the improvement of agricultural practices [15].

The age of the farmers is another construct considered in the studies [12–17]. Older farmers are less likely to adopt modern agricultural practices [15].

Wu [10] concludes that the perceived performance expectation of ICT for farmers will have a positive outcome on behavioral intention. Social influence on information technology for farmers will have a positive impact on behavioral intention. Facilitation conditions and behavioral intention on information technology for farmers will have a positive impact on usage behavior. In conclusion, farmers will have a positive impact on adopting information technology services by the effort of expectations, performance expectations, social, and environmental factors.

The age of the farmer has a variable effect on the adoption decision. In some cases, younger age was recognized as relevant for adoption, as it possibly confers

Table 1 Constructs identified in the studies

Constructs	Author
Digital education/education and knowledge/level of ICT education/education/level of education	[7, 7, 12–17]
ICT availability/access to ICT/facilitating conditions/Internet access	[7, 8, 10, 15, 17–19]
Age of farmers	[12–17]
Subjective norm/social influence	[10, 18–22]
Farm size	[3, 13, 14, 16, 17]
Perceived usefulness	[3, 20–23]
Attitude	[3, 14, 20, 22]
Perceived effort expectancy	[10, 19, 20, 23]
Perceived economic well-being/economic income	[3, 16, 20]
Factor in farmers´ capacity and skills/farmer´s ICT skills	[3, 7, 15]
Behavioral intention/motivation to use	[3, 10, 18]
Use of ICT	[8, 24]
Farm characteristics	[13, 16]
Culture (social, organizational)	[3, 24]
Role of government/government support	[3, 24]
Perceived performance expectancy	[10, 19]
ICT costs	[3, 25]

longer work horizons. The expectation of perceived effort in information technology for farmers will have a positive impact on behavioral intention.

In another study [3], the factors were classified into two categories such as internal factors and external factors. The factors farmer's attitude toward the use of IT, training and education, and cost of ICT was the most scored among the internal factors. Among the external factors, government support is mentioned as the most important. Facilitating condition has been considered the most influential factor after effort expectancy [19].

Other factors have been mentioned such as personal characteristics, access to electricity, number of crops grown, or dissemination of information (see Table 2).

The adoption models more used are TAM [20, 22–24] and UTAUT [10, 19, 20]. Other studies were based on the theory TPB [18, 22] or the combination of TPB and TAM [22].

This research has limitations. One of the limitations of this research is that the number of articles is low, and it is necessary to increase this number in future using the technique of systematic review.

Table 2 Other constructs identified in the studies

Constructs	Author
Computer self-efficacy	[23]
Personal characteristics	[14]
Access to electricity	[8]
Perceived behavior control	[18]
Innovativeness	[13]
Impact of social networks	[24]
Dissemination of information	[24]
Social category	[16]
Number of crops grown	[16]
Commercial orientation	[16]
Access to mobile phone	[17]
Access to radio or television	[17]
Perceived utility	[20]
Computer anxiety	[23]
Relative advantage	[25]
Perceived complexity	[25]
Employee capability	[25]
Top management support	[25]
Perceived vendor support	[25]

5 Conclusions

Information and communication technologies (ICT) have been incorporated into different sectors, including the agricultural sector. The applications are diverse: price prediction of agricultural products, e-commerce, smart cell phones, information systems, among others. The adoption of ICTs is conditioned by a number of factors. A total of 17 articles were analyzed to determine the constructs or factors that condition the adoption of ICTs in the agricultural sector. The results show that the factors digital education, availability and access to ICT, farmers' age, subjective norm, farm size, perceived usefulness, and attitude toward ICT use are the most frequent. Knowing the constructs will allow us to propose a model for ICT adoption in agricultural associations in future. It is necessary to invest in technological infrastructure, in policies aimed at a digital agenda that includes expanding access to ICTs for the population in general, especially the rural population, which is estimated to be the least served and where there is a great demand for ICTs.

References

1. Sabu K, Kumar T (2020) Predictive analytics in agriculture: forecasting prices of Arecanuts in Kerala. Procedia Comput Sci 171:699–708
2. Zhang D, Zang G, Li J, Ma K, Liu H (2018) Prediction of soybean price in China using QR RBF neural network model. Comput Electron Agric 154:10–17
3. Wongsim M, Sonthiprasat R, Surinta O (2018) Factors influencing the adoption of agricultural management information systems in Thailand. In: 2018 3rd Technology innovation management and engineering science international conference (TIMES-iCON) on proceedings, pp 1–4
4. Davis F (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q Manag Inf Syst 13:319–339
5. Venkatesh V, Morris M, Davis G, Davis F (2003) User acceptance of information technology: Toward a unified view. MIS Q 425–478
6. Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50(2):179–211
7. Pivoto D, Waquil P, Talamini E, Finocchio C, Dalla Corte V, de Vargas Mores G (2018) Scientific development of smart farming technologies and their application in Brazil. Inform Process Agric 5(1):21–32
8. Perez-Estebanez R, Urquía-Grande E, Rautiainen A (2018) Technological and economic factors determining ICT level: evidence from rural micro-businesses in democratic Republic of Congo. J Int Dev 30(1):118–133
9. Gestión (2018) https://gestion.pe/economia/brecha-productividad-agricultura-aumenta-peru-dice-banco-mundial-228415-noticia/
10. Wu L (2012) An empirical research on poor rural agricultural information technology services to adopt. Procedia Eng 29:1578–1583
11. van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric map- ping. Scientometrics 84(2):523–538
12. Botsiou M, Dagdilelis V (2012) Aspects of incorporation of ICT in the Greek agricultural enterprises: the case of a prefecture. Procedia Technol 8:387–396
13. Michels M, Fecke W, Feil JH, Musshoff O, Pigisch J, Krone S (2020) Smartphone adoption and use in agriculture: empirical evidence from Germany. Precision Agric 21(2):403–425

14. Taragola N, Van Lierde D (2010) Factors affecting the Internet behaviour of horticultural growers in Flanders Belgium. Compute Electron Agric 70(2):369–379
15. Wims P (2011) ICTs in Irish agriculture: can ICTs improve communication between Agribusiness and Farmers? In: HAICTA, pp 103–119
16. Ali J (2012) Factors affecting the adoption of information and communication technologies (ICTs) for farming decisions. J Agric Food Inform 13(1):78–96
17. Mittal S, Mehar M (2015) Socio-economic factors affecting adoption of modern information and communication technology by farmers in India: Analysis using multivariate probit model. J Agric Educ Ext 22(2):199–212
18. Zamani-Miandashti N, Memarbashi P, Khalighzadeh P (2013) The prediction of internet utilization behavior of undergraduate agricultural students: an application of the theory of planned behavior. Int Inform Library Rev 45(3–4):114–126
19. Wichean A, Sungsanit M (2022) Factors influencing the intentions to adopt technology of the broiler farmer in livestock region 3 Thailand. Trends Sci 19(1):1707–1707
20. Verma P, Sinha N (2018) Integrating perceived economic wellbeing to technology acceptance model: the case of mobile based agricultural extension service. Technol Forecast Soc Chang 126:207–216
21. Kabir KH, Hassan F, Mukta MZN, Roy D, Darr D, Leggette H, Ullah S (2022) Application of the technology acceptance model to assess the use and preferences of ICTs among field-level extension officers in Bangladesh. Digital Geogr Soc 100027
22. Ulhaq I, Pham N, Le V, Pham H, Le T (2022) Factors influencing intention to adopt ICT among intensive shrimp farmers. Aquaculture 547:737407
23. Zarafshani K, Solaymani A, D'Itri M, Helms MM, Sanjabi S (2020) Evaluating technology acceptance in agricultural education in Iran: a study of vocational agriculture teachers. Soc Sci Humanit Open 2(1):100041
24. Aleke B, Ojiako U, Wainwright D (2011) ICT adoption in developing countries: perspectives from small-scale agribusinesses. J Enterp Inf Manag 24(1):68–84
25. Abdullahi H, Hassan A, Mahmud M, Ali A (2021) Determinants of ICT adoption among small scale agribusiness enterprises in Somalia. Int J Eng Trends Technol 69(2):68–76

Effective Biometric Technology Used with Big Data

Abdulrahman S. Alenizi and Khamis A. Al-Karawi

Abstract The research has proposed a detailed study into the perceptive of government employees about the entry of the concept of biometric authentication like voice recognition at the respective workplaces in the State of Kuwait. Studies were carried out on employees regarding the factors that influence the employees to adopt any new technology, which helps in improved reception of biometric technology in the Big data applications. To derive the required data for this study, a mix of interviews and questionnaires was done; managers had to give the interviews, and employees had to complete questionnaires which were provided by four particular government organizations in the State of Kuwait to understand the perception of the employees using voice recognition biometrics. It was seen from the study that there was a notable cultural and digital space between the required authentication solutions issued by the management and the technological knowledge of the employees. Due to the situation of misuse, and mainly a lack of trust in the newly introduced technology and different management intentions, managers feel more responsible for minimizing the prevalent gaps. In the context of voice recognition implementation, it was more essential to tackle the employees' resistance. It is highly recommended to have proper orientation and knowledge in voice biometrics much before introducing the technology into the organization.

Keywords Big data · Voice recognition · Biometric · Users' perceptions

1 Introduction

The evolution of new dimensions is constant in our lives with every new technology that emerges. This helps to create interactions amongst users and the respective governments via the help of electronic services. Governments seek the effective and

A. S. Alenizi (✉)
Business School, University of Salford, Manchester, UK
e-mail: a.alenizi@edu.salford.ac.uk

K. A. Al-Karawi
University of Diyala, Baqubah, Diyala, Iraq

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_22

efficient use of technology to deliver their services electronically with more ease [1–3]. Hence, the electronic government has emerged to become quite a vital application worldwide. For the function of Big data applications, users need to submit their data to ensure an effective, efficient, safe technology, especially for providing reliable methods for users and secure data systems. With the introduction of a Big data system, the government has to tackle issues like user authentication, privacy, and information security, where biometric would prove to be quite an ideal solution to tackle the related concerns [4, 5]. It offers the proper identification facility of individuals and the provision to control and safeguard the integrity of subtle, secure data that would be saved in information systems [6, 7]. Thus, many governments have taken up the use of biometric authentication systems for efficient functioning and delivery of services. Simultaneously, this adoption process of fingerprint biometrics into Big data has become a vital element for political planning in several governments. User acceptance is considered essential for biometrics to be successfully implemented [2, 8]. Users may possess a straight influence on the functioning of biometric systems. Hence, their requirements must be considered carefully, irrespective of how vague or rough they may be [9, 10]. The aim of this research is to understand the factors that impact the employees' welcoming of the new biometrics concept and explain to them how to do the voice recognition biometric implementation in Big data applications.

2 Background

2.1 Big Data Challenges

Although the governments have made many efforts in this regard, the success of the Big data programme depends on the users' readiness to utilize and receive the Big data digital services using new Information and Communication Technology (ICT) [4, 11]. Remarkably, security is one of the critical reasons. Extensive data systems are reluctant because of ongoing perceived security issues [12, 13]. Table 1 shows the details of the technological challenges.

2.2 Digital and Cultural Gap

Digital division applies to the difference between the very standard and high technology open community of citizens [14, 15]. That may be due to several factors, such as the shortage of funding, high school, and electronic literacy. However, the digital divide challenges the effectiveness of free digital apps. For example, citizens in rural areas and neighbourhoods in the city's inner-city have less Internet connectivity than others, whilst others who have never had devices refuse to utilize modern technologies [7, 16]. There may be a digital gap in the Kuwaiti situation due to a

Table 1 Shows the details of the significant data challenges

Challenges	Example and of challenges
Technological	
• I.T. infrastructure	– A lack of capacity for networking – Inadequate framework integration – Hardware and applications upgraded poorly – Compatibility and sophistication of the structures in operation
• Security	– Failure to secure transactions – Trust in digital services online and government – Failure of public sector security hardware
• Availability	– Inability on request to offer resources and details – React steadily and infringe on the demands of people
• Accessibility	– Difficulty by individuals with disabilities to navigate the scheme – The scope of the Internet is low
• Website design	– Two limited languages for the contents of the website – The simplicity of usage perceived and perceived utility
User-aspects	
• Lack of awareness • ICT skills	– Failure to inform about and learn from modern open government programmes – Inability to support extensive data orientation campaigns – I.T. skills gaps amongst digital available administration users; (i.e. citizens, employees, I.T. staff) – Lack of essential awareness of healthy web activities
Social	
• Culture	– The religion of the traditional and tribal – The trouble with language – Resistance to transition
Financial	
• Lack of budget/high cost	– There high operating and repair costs – High-security systems expenditure High

lack of technological expertise and practise. The results segment shows digital and cultural gaps between government employees' technical knowledge and the growing need to cope with modern technologies.

2.3 Identification and Verification Technology

Biometrics refers to identifying the users automatically depending upon the person's particular behavioural or physiological traits [17, 18]. A biometric system, which is a pattern recognition system, works via deriving an individual's biometric information, obtaining a feature set, which is then compared with the already saved template data kept in the database. The biometric systems are working in two modes: either in the verification or in the identification mode, depending on the context. The verification

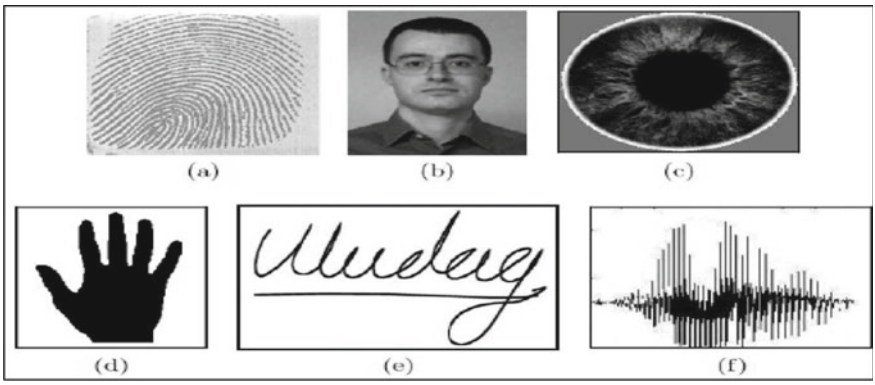


Fig. 1 Sample biometric traits: **a** fingerprint, **b** face, **c** iris, **d** hand geometry, **e** signature, and **f** voice

mode refers to the method where the system tests the authenticity of a person's identity by comparing it with the acquired biometric data with the biometric template(s) kept in the system's database [5, 10, 19]. The verification process is intended to avoid any situation where several people may use the same identity. On the other hand, identification mode is a negative recognition that prevents a person from using numerous identities. In this work, voice recognition is used as a biometric. A sample of biometrics can be seen in Fig. 1, used in verification or identification mode, the voice recognition biometric involved in this work.

2.4 Voice Recognition System

Automatic speaker recognition using unique voice details is found in speech waves by recognizing an unnamed speaker from multiple speakers. The aim of speaking person recognition is to isolate, classify, and use this speaking knowledge for identifying purposes. It differs from the recognition of expression and language knowledge since such definitions are about language recognition and voice interpretation [15, 20]. Speaker recognition comprises feature extraction, classification, design matching, decision-making processor modules, and speaker modelling. Voice recognition intends to derive speaker-specific data, illustrate them, and then apply them for identification purposes. In the testing phase, a sample of speech of the users is matched with the previously stored speaker models, which were structured in the training phase. The speaker's award may be split into the verification and identification of the speaker [13, 15]. Speaker tests are established to determine if the speaker claims to be a speaker or not. The identity of the speaker may often be split into two groups of problems: the closed one and the open one. Alternatively, the individual who speaks a test utterance will determine if they belong to a community of recognized PIN codes [7, 12, 21].

2.5 Some of Biometric Used in Big Data Applications

With the use of voice recognition technology, the Big data and e-government intend to offer better services to its citizens with secure and efficient access to data information, offering identification of reliable individuals along with the facility for protecting and controlling the honour of delicate data information systems. Scholars like Ashbourne [9], Bonsor and Johnson [22, 23], and Scott [2] state that biometric technology can be widely used for Big data and e-government schemes. Biometric technology is presently applied in banking transactions, security access, and e-voting to guarantee that voters are not voting more than once. With the help of biometric technology, it has become possible for governments to minimize any fraud during any transactions or through elections. It is, however, confirmed that biometric offers a significant amount of authentication by recognizing people, yet there are negative features too [24].

2.6 Concerns About the Use of Biometric Technology

Biometrics may offer a high degree of security by recognizing individuals through their physiological and behavioural traits, but they will also fail to provide biometrics. For example, fingerprints left on certain appliances can be saturated, faint, or challenging to manage, mainly when the skin is wet or dry. Hand identification may often be unsuccessful whilst the hand is impaired because no outcome matches the photographs in the database already [24]. In addition, there are many obstacles to the predecessor recognition scheme, such as sound noise and reverberation, which may impair the accuracy of recognition performance [25] (Fig. 2).

3 Methodology

Analysing the literature on the application of biometrics to voice recognition led to finding the available methods for an exploratory study related to the same. Two questions were framed to derive descriptive information from both potential and current users of the voice recognition application. The research was structured to obtain answers to the questions below;

- What are the manager's conceptions concerning utilizing voice recognition in Big data applications?
- What are the employees' conceptions concerning utilizing voice recognition in Big data applications?

The study was approved in three stages with the three main groups of people, top and middle managers, and employees, and the sampling method was purposive. This

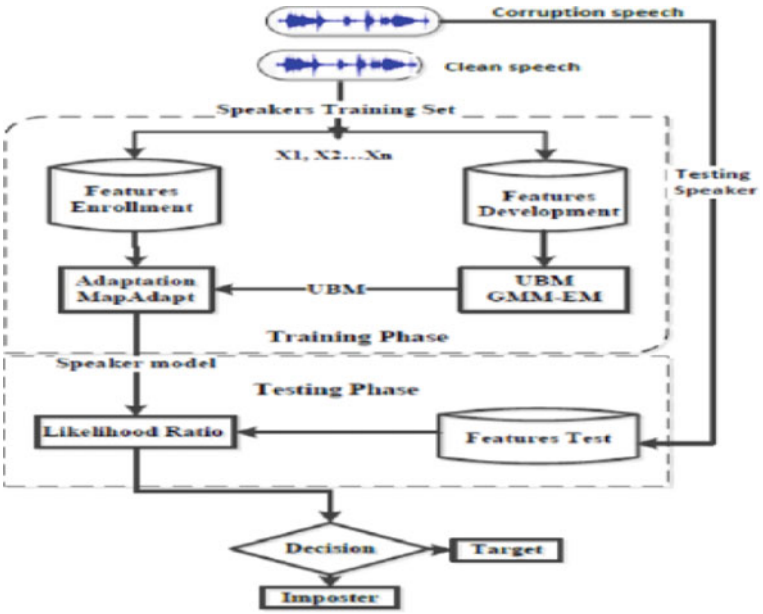


Fig. 2 Framework of the speaker recognition system

particular sampling process [26] is a pattern where "specific settings, people, or activities are nominated purposely to give data that cannot be acquired as well from other selections. The questionnaire was formed based on user acceptance of voice recognition tools. Questions were discussed more in detail to understand the elements that impact the use of the biometric application, like voice recognition. A quantitative method provides a helpful trend analysis. Managers and employees had particular questionnaires to understand their perspectives about the biometric application. The usage of several or hybrid approaches both in quality and quantitative ways thereby balances the limitation of one approach by the strengths of the other [27].

3.1 Data Gathering and Examination

The data relating to this study was derived from interviews and questionnaires. The maximum interview time was estimated at 40 min, but the actual recording duration was about 20–30 min. A questionnaire pattern was used with 60 employees from different organizations in Kuwait [28]. Furthermore, interviews were conducted in Kuwait using Skype software with ten managers of the Vocational education and training and National Bank of Kuwait. To obtain personally meaningful information from the participants, open-ended questions were used for the interviews [27]. Both

answers were preserved in the Statistical Social Science Package (SPSS) programme used for research.

3.2 The Finding

In the past, in most agencies of Kuwait, the recording of manual signature used to be the procedure for confirming the employees' attendance. This process, however, has some negative points like, for instance, the employees could sign in for anyone else and maybe even not record the correct time of signing. It was realized that this is not an efficient and effective procedure for recording attendance which paved the way for introducing voice recognition technology.

3.3 Interview Results of Managers

Likert scale form containing questions ordered from (1–5) was used in this interview. Analysis in the question-by-question pattern is shown in this portion. Question1 represents the lower rank level, and level five is the highest rank. An analysis in the pattern of question is obtainable in the following.

3.3.1 What Cultural Gap Do You Observe Amongst the Employees' Level of Technological Knowledge and the Level of Voice Recognition Utilized?

This question aims to investigate the cultural space amongst the degree of voice recognition technology and technological experience of the employees in the respective organization.

- Three of the candidates credited the prevalent technical, cultural space amidst the staff to the standards of technological literateness.
- Two of the candidates claimed the prevalent technical, cultural space related to the employees' age, meaning the gap is more comprehensive as the employee grows older.
- According to one of the candidates, the cultural gap was the result of the perception that using this technology showed the management's mistrust in the employees, persuading them against the application of this technology
- However, the two candidates did not mainly attribute the cultural space to any exact reason.
- As two candidates, they did not even find any cultural gap.

3.3.2 Do You Agree on a Level of Responsibility for Tapering This Cultural Gap?

This question checks the top and middle managers' estimated accountability for minimizing the cultural gap between their employees and the levels of voice recognition technology available at their workplace.

- Sixteen amongst the candidates, five considered them responsible for minimizing the cultural gap. They initiated steps and processes to encourage employees about the need for awareness of technology and its utilities.
- Four individuals did not deliberate themselves as responsible for minimizing the cultural gap.

3.3.3 Have You Known Any Problems in Dealing with This Technology? if Yes, What Were They?

The interviewees brought forward 10 responses, set out in the categories of difficulty or frustration that they had faced, as well as what they thought were some of the most common difficulties encountered.

- The resistance of the employee (10 respondents);
- Breaking down and disabling the voice recognition instrument by certain employees (5respondents);
- Failure of the system (4 respondents);
- System failing to identify certain users (5 respondents).

3.3.4 What Are the Critical Barriers of Using Voice Recognition in Your Organization?

This question points to the top and middle managers' perspectives regarding the obstacles to implementing voice recognition technology in their organizations.

- The answers to this problem were associated with technological and digital culture and the struggle to change for which the employees validated right at the start of the arrangement.

3.3.5 How Do You Think Utilizing Voice Recognition Innovation Influences the Self-Perceived Social Level of Your Employee?

This question looks at the social influence created by the implementation of voice recognition technology on their employees and society.

- According to six respondents, more positive effects could be seen with stricter regulations.
- Four of the respondents pointed out the negative features of this technology.

- Optimistic and passive impacts were known to be the cause of the following:
- Four of the respondents checked the suspicion worries felt by some employees.
- One of the individuals testified that he felt proud of using technology, whilst his friends who do not use it do not feel the same.

There is a cultural and digital gap validated with the support of the technical awareness of employees and also the expected verification results ensured by the management.

3.4 Questionnaire Results of Government Employees

Likert scale form containing questions ordered from (1–5) used in this survey. Question1 represents the lower level of importance, and level five is the highest importance. Survey was spread amongst 60 employees working at a government organization.

3.4.1 How Do You Think Using Voice Recognition Technology in a Government Organization is Significant?

The following mentioned are the derived responses:

- None of the concerned participants thought it was unimportant.
- 28.6% consider it essential.
- 8.7% view it as quite important.
- A minority (40.3%) rate of participants are unaware of the value of applying the system of voice recognition technology at their respective workplace

3.4.2 Do You Think It is Essential to Be Aware of This Technique Before it is Used?

The notion of awareness contains the elements of employees' information, statement, and instruction. The respondents differentiated the degree of significance in the following ways—1) What is the role of technology in this area?

- Only 3% of the participants felt there is no need to enhance the employees' alertness about the related biometric before its installation
- 13.5% consider it essential.
- Most of the respondents (48.46%) have assumed that it is essential to have know-how before using speaker recognition biometrics.

3.4.3 Do You Think Utilizing This Biometric in Your Work Means Employers Suspect Employees?

The response to this question analyses the users of the view of the employer suspicion that comes from the introduction and use of voice recognition technology.

- 32.6% of respondents declared that it does not intend mistrust.
- 9.7% consider its distrust.
- 19.9% are sure that it is mistrust.
- 27.8% of participants are not sure if it is mistrust or not.

4 Discussion

The derived outcomes show that after the survey, 90% of the managers have agreed that there is a cultural/digital gap as the result of less familiarity of the employees with the technology and the organization's embracing of voice recognition biometrics. At the same time, 45% of other respondents stressed the negative effect of using this biometric, which was based on the findings by Coventry [28], who focused on the feebleness of the cultural and social consciousness of the ideas and DOG applications. This strengthens the trials in using Big data in Kuwait, which denotes a lack of awareness classes about the application of Big data, inadequate curricula for official education for preparing for the information age, and the citizens' illiteracy of computer knowledge. 34.8% of the survey respondents are not sure if the arrival of this new technology denotes mistrust. 23.6% of respondents are confident that using this technology indicates that employers lack trust in their employees. This could be due to factors like absence of knowledge via consultation, information, notification, and extent of computer literacy.

5 Conclusion Remark

Research learning was conducted to understand the perceptions of government employees regarding factors linked with the arrival of voice recognition biometrics in the workplace. There must be an orientation and awareness programme about voice recognition before introducing this technology into the organization. Top and middle managers pointed out the employees' resistance to installing the technology before even installing it. Most managers refused to think of themselves as responsible for minimizing the cultural and digital gap concerning voice recognition technology.

Acknowledgements The authors would like to express their gratitude to Kuwait's Public Authority for Applied Education and Training for their assistance with this project. The editors and anonymous

reviewers deserve thanks for their insightful comments and suggestions, which helped to improve the paper.

References

1. AlAwadhi SA (2007) E-government in Kuwait: attitudes and perceptions. Loughborough University
2. Acton T, Hughes M, Scott M (2005) An assessment of biometric identities as a standard for e-government services
3. Alenizi AS, Al-karawi KA (2022) Cloud computing adoption-based digital open government services: challenges and barriers. In: Proceedings of sixth international congress on information and communication technology. Springer
4. Dearstyne BW (2001) E-business, e-government & information proficiency(cover story). Inf Manag J 35(4):16–23
5. Al-Karawi KA, Ahmed ST (2021) Model selection toward robustness speaker verification in reverberant conditions. Multimedia Tools App 1–18
6. McLindin BA (2005) Improving the performance of two dimensional facial recognition systems: the development of a generic model for biometric technology variables in operational environments
7. Al-Karawi KA (2020) Mitigate the reverberation effect on the speaker verification performance using different methods. Int J Speech Technol 1–11
8. Giesing I (2003) User response to biometric. University of Pretoria, Pretoria, South Africa, pp 95–135
9. Ashbourn J (2015) Practical biometrics: From aspiration to implementation. Springer
10. Al-Karawi K (2018) Robust speaker recognition in reverberant condition-toward greater biometric security. University of Salford
11. Mohammed DY et al (2020) Mitigate the Reverberant Effects on Speaker Recognition via Multi-training. Springer International Publishing, Cham
12. Chen K, Xie D, Chi H (1996) Speaker identification using time-delay HMEs. Int J Neural Syst 7(01):29–43
13. Al-Karawi KA, Mohammed DY (2021) Improving short utterance speaker verification by combining MFCC and Entrocy in Noisy conditions. Multimedia Tools Appl 80(14):22231–22249
14. Blau A (2002) Access isn't enough: merely connecting people and computers won't close the digital divide. Am Libr 33(6):50–52
15. Al-Karawi KA et al (2015) Automatic speaker recognition system in adverse conditions-implication of noise and reverberation on system performance. Int J Inf Electron Eng 5(6):423
16. Alharbi SJ (2006) Perceptions of faculty and students toward the obstacles of implementing e-government in educational institutions in Saudi Arabia
17. Chen W-S et al (2005) Personal identification technique based on human iris recognition with wavelet transform. In: Proceedings of (ICASSP'05) IEEE international conference on acoustics, speech, and signal processing 2005. IEEE
18. Al-Karawi KA, Mohammed DY (2019) Early reflection detection using autocorrelation to improve robustness of speaker verification in reverberant conditions. Int J Speech Technol 1–8
19. Reynolds DA (2002) An overview of automatic speaker recognition technology. In: 2002 IEEE international conference on acoustics, speech, and signal processing (ICASSP)
20. Gish H, Schmidt M (1994) Text-independent speaker identification. IEEE Signal Process Mag 11(4):18–32

21. Al-Karawi KA, Mohammed DY (2019) Early reflection detection using autocorrelation to improve robustness of speaker verification in reverberant conditions. Int J Speech Technol 22(4):1077–1084
22. Gudavalli M, Kumar DS, Raju SV (2014) Securing e-governance services through biometrics. Int J Secur Appl 8(1):103–112
23. Al-Karawi KA, Li F (2017) Robust speaker verification in reverberant conditions using estimated acoustic parameters—a maximum likelihood estimation and training on the fly approach. In: 2017 Seventh international conference on innovative computing technology (INTECH). IEEE
24. Bolle RM et al (2013) Guide to biometrics. Springer Science & Business Media
25. Zhao X, Wang Y, Wang D (2014) Robust speaker identification in noisy and reverberant conditions
26. Maxwell JA (2012) Qualitative research design: an interactive approach. vol 41. Sage publications
27. Murray A, Pace R, Scott D (2004) Research: a commonsense approach. Thomson Social Science Press, Victoria
28. Frankel J, Wallen N (2000) How to evaluate & design research in education. McGraw-Hill, Boston, MA

Dropout in Higher Education and Determinant Factors

Sussy Bayona-Oré

Abstract One crucial topic in higher education that have negative consequences is student dropout. In academic organizations there is a duality between retention and dropout. This situation is not only detrimental to the student who sees his or her dreams of completing a professional career cut short, but also to the educational organizations. Knowing the factors that cause the student dropout allows educational organizations to take actions to prevent dropout. The results of the literature review on factors related to dropout, as well as the categories used by the authors to categorize them, are presented in this article. The results show that institutional, academic, individual, economic, and vocational are the categories most used by authors. Knowing the determining factors will make it possible to analyze how these factors influence dropout.

Keywords Dropout · Higher education · Factors · Categories · Student dropout

1 Introduction

Nowadays many students who finish their secondary studies choose to pursue a university career. This is a new experience and generates many expectations. However, the student's dream of completing the university career may be truncated. Admitting to university, in addition to being a change that implies responsibility and independence of the student, requires the ability to adapt to a new study environment. Also, the level of preparation of the student [1] and motivation for the study program [2].

The decision made by the student to dropout of their studies is a phenomenon that has been considered a social problem with negative consequences. Student dropout is a common problem in educational institutions. Not completing this educational process places the person at a disadvantage for the rest of his or her life [3]. For that reason, evaluating the student dropout rate becomes a vital concern [4].

S. Bayona-Oré (✉)
Dirección de Investigación, Universidad Autónoma del Perú, Av. Panamericana Sur Km. 16.3, Villa EL Salvador, Lima, Perú
e-mail: sbayonao@hotmail.com

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_23

The above is the impact at the individual level, but dropout has a socio-economic impact and puts young people at risk [5]. Dropout is considered a complex problem, which has motivated interest in conducting studies on the factors that condition it.

Leaving a study program prematurely, before achieving the degree [6] is called student dropout. Dropout is considered a topic of interest in the educational field [7] and has given rise to different proposals to explain it. Tinto's model [8] is the first and most recognized model of university dropout [9], which conceives academic persistence as a phenomenon that depends on various factors, i.e., it is multifactorial, and is favored by the integration between the student and the academic environment.

The impact of student dropout affects educational institutions and there is an impact at the individual (student), economic and social levels [10]. In some cases, despite the various initiatives implemented by universities to retain the student [11] the student may decide to abandon studies.

There are several factors that condition student dropout. That is, the permanence of the student is conditioned by individual, institutional, economic factors [12], physiological or related variables [9]. This makes dropout a multidimensional phenomenon, due to the interaction of different factors.

The purpose of this article is to determine the most common critical factors that have been mentioned in the literature. Knowing the factors that condition dropout makes it possible to develop strategies to reduce the dropout rate. The discussion of determinants has been a topic of interest and debate by researchers and educational policy makers [13]. Knowing the determinants allows policymakers to take preventive actions before corrective ones. This article is a contribution to knowledge and serves as a basis for future research.

Therefore, after the introduction, the rest of the paper is structured as follows: Sect. 2 presents the background of the research work; Sect. 3 presents the methodology used; Sect. 4 presents the results. Finally, conclusions and future work are included in Sect. 5.

2 Background

The interest in studying student dropout is because various factors interact with each other to predict it [14]. Student dropout affects students because they stop graduating. Dropout is a process of social and academic interaction and integration of the student [15].

The previously mentioned is related to personal and institutional objectives. Likewise, the student's own factors are involved in dropout, as well as the student's ability to integrate into a new academic environment to achieve his or her goals. It is necessary to differentiate between two types of dropout voluntary when the student voluntarily decides to leave the career, and involuntary when this decision is institutional, either because of performance or discipline.

There is not a standard definition for dropout. Tinto [8] define "dropout as student who enroll in college but do not reenroll or do not complete their intended degree program or set of courses".

Other definition is "the premature abandonment of a study program before reaching the degree and considers a sufficiently long time to rule out the possibility of the student rejoining" [6]. Dropout is defined "those students who do not return to the college in which they enrolled, have no definite plans to return, and do not transfer to another institution of higher education" [16].

2.1 Determining Factors and Dropout

The causes that influence student dropout are diverse and have been studied to date. Moreover, student dropout has increased due to the impact of COVID-19. The reasons for dropout have been discussed by various authors who argue that there are several reasons why students fail in universities.

Dropout is due to institutional, academic, economic, academic, vocational characteristics. Also, in the first four academic cycles there may be higher dropout rates. For that reason, an important indicator in the accreditation system is dropout, so universities should have mechanisms to identify problems in the student's progress, have a tutoring system or reinforcement and leveling programs.

Student dropout can occur (1) in students who entered and left the educational system without finishing it and (2) in those who finish the course, but do not continue in the next one [17]. Likewise, the first semesters of study are conducive to dropout and is associated with economic factors [17]. In other cases, the benefits brought by information and communication technologies (ICT) can be overshadowed as a cause of the decision to dropout of college. Mostly, when ICTs are used inappropriately.

Bher et al. [7] conducted a study with the participation of students who had dropped out and found that lack of interest in the field of study and wrong expectations were the main causes of dropout. In another study [18] the most influential factor was the time gap between secondary and tertiary education.

Morelli et al. [9] argue that a student who can manage university activities is less likely to dropout. Therefore, it is important that students learn to manage their time, know study methods, techniques to reduce stress, interpersonal skills, and techniques to improve concentration, among others.

Academic performance is considered as a determinant of student dropout. Goundar et al. [19] argue that some students in computer science programs consider it a difficult discipline and that students who do not do well in courses are more likely to dropout. In other words, academic performance is one of the main reasons for abandoning studies [20].

Zhao et al. [21] argue that the lack of commitment to science, technology, mathematics and engineering and the high dropout rates are related to the number of unfilled positions in the ICT area. Another factor mentioned is [22], especially when the experience is negative, it can generate stress, anxiety and lack of concentration.

In addition, in the process of deciding to continue the studies or not to continue, preceding the abandonment of the studies, mental health problems may arise.

Wild [2] found connections between motivation and student behavior. The reason is that as the study time progresses, the motivation for the program decreases.

A problem that students present in the first cycles is the adaptation to the university system, because it is a change, and in some cases, they do not know how to manage this change. In other cases, students may present low academic performance and decide to stop studying, and consequently young people cannot meet their personal and/or family goals, which is exacerbated when the student does not assume responsibility.

3 Methodology

To determine the determinants of dropout a literature review was conducted. For this purpose, scientific articles published in the period 2020-22 were reviewed.

Scopus was used to identify the scientific articles related to student dropout. The terms *dropout*, "*drop_out*", "*university students*" and "*higher education*" were used to create the search string. The term "*packet dropout*" was excluded.

The inclusion criteria were scientific articles that include the classification of factors related to student dropout. Short articles were not considered. To extract the relevant information from each of the scientific articles, an Excel format was designed with the following information: general data of the article, categories used to classify the factors related to dropout, factors and definitions of the factors.

4 Results

There is a set of factors that determine the student dropout. Various authors propose categories to group the factors. Next, the results of the review of scientific articles are presented with the purpose of identifying the factors and the categories used to classify them.

4.1 Categories Used to Classify the Factors

The categories used to classify the factors related to dropout are: (1) *individual* related to students [12, 23, 24] such as academic information and personal information [25] or *performance* [26], (2) *economic* [12, 23, 27], (3) *academic* [23, 24, 27], (4) *organizational* [9] or *institutional* related to university [12, 24, 26, 28, 29], (5) *external factors* related to information from the university, the environment and student support [25], (6) *related variable* [9], and (7) *psychological* [9].

Other categories such as (1) *environmental*, (2) *motivational* [27], (3) *vocational* [26, 27] and *socio-economic* [24] are mentioned. Additionally, Ambiel et al. [26] propose the categories (1) *support* that includes the financial difficulties or family or work conflicts, (2) *interpersonal* or the relationship with their peers, and (3) *career* related to future professional prospects.

Chopra and Menon [28] propose other categories such as (1) *family oriented*, (2) *community oriented*, and (3) *social* media oriented. Bussu et al. [29] propose the categories (1) *individual characteristics* such as gender and age, (2) *student background*, such us family income and education, (3) *institutional environment* such as teaching and research quality, and (4) *student satisfaction*.

The categories used by the authors to classify dropout factors are presented in Table 1.

Analysis of the results shows that the authors classify the factors according to their criteria and that there is no standardization. It is also observed that the same factor can be considered in two categories. Among the categories most used by the authors of the selected studies are institutional, academic, individual, economic, and vocational.

Table 1 Categories used to classify the factors and student dropout

Author	Categories
Morelli et al. [9]	Psychological, organizational or related to the university, and related variables
Aina et al. [12]	Individual or related to the students, institutional related with the university, and economic
Aldahmashi et al. [23]	Economic, individual, and academic
Guzmán et al. [24]	Individual, institutional, socio-economic, and academic
de Oliveira et al. [25]	Aspects related to the student such as academic information and personal information and external factors such as information from the university, the environment, and student support
Ambiel et al. [26]	Institutional, support, vocational, performance, interpersonal, and career
Rodríguez-Pineda et al. [27]	Academic, environmental, motivational, economic, and vocational
Chopra and Menon [28]	Student oriented, institution oriented, family oriented, community oriented, and social media oriented
Bussu et al. [29]	Individual characteristics, student background, institutional environment, and student satisfaction

4.2 Factors Related to Dropout Student

Some factors mentioned in the *Institutional* factors are information about the university, the environment, the support provided to students, teacher-student interaction, academic support or institutional commitment.

Some factors mentioned in the *Economic* category are: students who work to continue their studies and lose their jobs or their salary is cut, the need to work to support the family's livelihood, accumulating student debts and not being able to pay them, the cost of university studies, the student's need to work to pay for housing, food or study expenses, and the student's need to work in order to pay the costs of living, food or study.

Some factors mentioned in the *Individuals* category are age, gender, student's ability to integrate into the academic system, social integration, motivation to continue studying, student behavior, level of commitment, student engagement, student's characteristics, poor cognitive ability, limited learning abilities, student's pregnancy or decision to become a parent, health problems, integration problems, interpersonal relationships, self-efficacy, unfulfilled expectations or emotional crisis.

The factors related to *Academic* category are not performed well in courses, gap between secondary and higher education or level of preparation of the student, academic information, academic performance, duration of the career, academic help from peers, support provide to students, lack of orientation, level of satisfaction, curriculum expectations or teaching methodology.

Some factors related to *Vocational* category are lack of interest and perception of little usefulness of the selected career, choice of course, perceived interest in the program, and permanence due to professional interest.

Other factors related to *Family oriented* are family culture, socio-economic status, parental support, family composition, lack of parent education, change of place of residence, change of marital status, illness of a family member requiring care and dysfunctional home.

Unlike other years, in 2021 the change from face-to-face classes to virtual classes due to COVID-19 pandemic the factor student-perceived quality of received education and social interaction are mentioned. Also, case of burnout and health integration [30].

5 Conclusions

Student dropout is a problem that affects universities in the international and national context. Several studies on dropout have been conducted and it continues to be a topic of interest to the academic community. Recently, with the impact of the COVID-19 pandemic, student dropout has become more important due to the concern of the education sector at the level of government, universities, teachers, students, parents, and others. In this research, a literature review was conducted to identify

the categories used to classify the factors and conditioning factors for dropping out of a study program. The categories more used to group the factors related to student dropout are institutional, academic, individual, economic, and vocational. Knowing the determining factors will make it possible to analyze how these factors influence dropout. As future work, the literature review will be expanded, applying the systematic review to then conduct a study to analyze which factors have the greatest influence on student dropout.

References

1. Aina C, Mussida C, Lombardi G (2022) Are business and economics alike? Quaderno DISCE 147
2. Wild S (2022) Trajectories of subject-interests development and influence factors in higher education. Curr Psychol 1–17. https://doi.org/10.1007/s12144-021-02691-7
3. OECD (2010) Overcoming school failure: policies that work. Project Description. https://www.oecd.org/education/school/45171670.pdf
4. OECD (2019) Education at a glance 2019. https://www.oecd-ilibrary.org/education/education-at-a-glance-2019_f8d7880d-en
5. Sharma C, Kukreja V, Kaur D (2022) Cause analysis of students' dropout rate using PSPP. In: Machine learning, advances in computing, renewable energy and communication, vol 768, pp 459–469
6. Himmel E (2002) Modelo de análisis de la deserción estudiantil en la educación superior. Calidad en la Educación 17:91–108
7. Behr A, Giese M, Teguim H, Theune K (2021) Motives for dropping out from higher education—an analysis of bachelor's degree students in Germany. Eur J Educ 56:325–343
8. Tinto V (1993) Leaving college: rethinking the causes and cures of student attrition, 2nd edn. University of Chicago Press, Chicago, IL
9. Morelli M, Chirumbolo A, Baiocco R, Cattelino E (2021) Academic failure: individual, organizational, and social factors. Educ Psychol 27(2):167–175
10. Kim D, Kim S (2018) Sustainable education: analyzing the determinants of university student dropout by nonlinear panel data models. Sustainability 10:1–18
11. Eather N, Mavilidi M, Sharp H, Parkes R (2022) Programmes targeting student retention/success and satisfaction/experience in higher education: a systematic review. J High Educ Policy Manag 44:1–39
12. Aina C, Baici E, Casalone G, Pastore F (2021) The determinants of university dropout: a review of the socio-economic literature. Socioecon Plann Sci 79:1–16
13. Mabel Z, Britton T (2018) Leaving late: understanding the extent and predictors of college late departure. Soc Sci Res 69:34–51
14. Bernardo A, Galve-González C, Núñez J, Almeida L (2022) A path model of university dropout predictors: the role of satisfaction, the use of self-regulation learning strategies and students' engagement. Sustainability 14(3):1–10
15. Tinto V (1975) Dropout from higher education: a theoretical synthesis of recent research. Rev Educ Res 45(1):89–125
16. Hoyt J, Winn B (2004) Understanding retention and college student bodies: differences between drop-outs, stop-outs, opt-outs, and transfer-outs. J Student Affairs Res Prac 41(3):395–417
17. Torres J, Acevedo D, Gallo L (2015) Causas y consecuencias de la deserción y repitencia escolar: una visión general en el contexto Latinoamericano. Cultura Educación y Sociedad 6(2):157–187
18. Berka P, Marek L (2021) Bachelor's degree student dropouts: who tend to stay and who tend to leave? Stud Educ Eval 70:1–13

19. Goundar S, Deb A, Lal G, Naseem M (2022) Using online student interactions to predict performance in a first-year computing science course. Technol Pedagogy Educ 1–19
20. Améstica-Rivas L, King-Domínguez A, Sanhueza D, Ramírez V (2021) Economic effects of dropout in university management: the case of a Chilean public university. Hallazgos 18(35):209–231
21. Zhao D, Muntean C, Chis A, Rozinaj G, Muntean G (2022) Game-based learning: enhancing student experience, knowledge gain, and usability in higher education programming courses. IEEE Trans Educ 1–12. https://doi.org/10.1109/TE.2021.3136914
22. Thabethe L, Mulondo S, Tugli A (2020) Psychological experiences of pregnant students at the University of Venda, Limpopo Province, South Africa. Afr J Reprod Health 24(3):18–23
23. Aldahmashi T, Algholaiqa T, Alrajhi Z, Althunayan T, Anjum I, y Almuqbil B (2021) Case-control study on personal and academic determinants of dropout among health profession students. Higher Educ Stud 11(2):120–126
24. Guzmán A, Barragán S, Cala F (2021) Rurality and dropout in virtual higher education programmes in Colombia. Sustainability 13:4953
25. de Oliveira C, Sobral S, Ferreira M, Moreira F (2021) How does learning analytics contribute to prevent students' dropout in higher education: a systematic literature review. Big Data Cogn Comput 5(4):64
26. Ambiel R, Cortez P, Salvador A (2021) Predição da Potencial Evasão Acadêmica entre Estudantes Trabalhadores e Não Trabalhadores. Psicologia: Teoria e Pesquisa 37:1–10
27. Rodríguez-Pineda M, Zamora-Araya J (2021) College student dropout: cohort study about possible causes. Uniciencia 35(1):19–37
28. Chopra A, Menon S (2021) Is job shadowing a panacea for educational drop outs? In: International conference on business and technology, vol 194, pp 975–987
29. Bussu A, Detotto C, Serra L (2020) Indicators to prevent university drop-out and delayed graduation: an Italian case. J Appl Res Higher Educ 12(2):230–249
30. Jacobo-Galicia G, Máynez Guaderrama AI, Cavazos-Arroyo J (2021) Fear of COVID, exhaustion and cynicism: its effect on the intention to drop out of college. Eur J Educ Psychol 4(1):1–18

A Text Classification for Vietnamese Feedback via PhoBERT-Based Deep Learning

Cu Vinh Loc, Truong Xuan Viet, Tran Hoang Viet, Le Hoang Thao, and Nguyen Hoang Viet

Abstract With the rapid development of social media platforms as well as the current pandemic, the majority of activities are performed online. The user comments obtained from the digital channels are crucial in order that the agencies or organizations can improve and develop their brand. Thus, an automatic system is necessary to analyze the sentiment of a customer feedback. Recently, the well-known pre-trained language models for Vietnamese (PhoBERT) have achieved high performance in comparison with other approaches. However, this method may not focus on the local information in the sentiment like phrases or fragments. In this paper, we propose a PhoBERT-based convolutional neural networks (CNN) for text classification. The output of contextualized embeddings of the PhoBERT's last four layers is fed into the CNN. This makes the network capable of obtaining more local information from the text. Besides, the PhoBERT output is also given to the transformer encoder layers in order to employ the self-attention technique, and this also makes the model more focused on the important information of the text segments. The experimental results demonstrate that the proposed approach gives competitive performance compared to the existing studies on three public datasets with Vietnamese texts.

Keywords Sentiment classification · PhoBERT · CNN · Deep learning · Text classification

C. V. Loc (✉) · T. X. Viet · T. H. Viet · L. H. Thao · N. H. Viet
Can Tho University, Can Tho, Vietnam
e-mail: cvloc@ctu.edu.vn

T. X. Viet
e-mail: txviet@ctu.edu.vn

T. H. Viet
e-mail: thviet@ctu.edu.vn

L. H. Thao
e-mail: lhthao@ctu.edu.vn

N. H. Viet
e-mail: nhviet@ctu.edu.vn

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_24

1 Introduction

Collecting customer feedback is a great way for agencies or organizations to understand the strengths and weaknesses of their products and services, and also to understand what the customers like and do not. This is also a way to quickly capture the psychology and need of customers to bring them the most perfect products and services. Nowadays, with the great development of science and technology, especially the explosion of transactions through the digital channels like social media and e-commerce has allowed people not only to share information on it, but also express their attitudes and views toward products, services and other social issues. Due to the current pandemic, besides, people are studying, working, and interacting from a distance. The online social networks are then continuously adapted and have generated a lot of opportunities for obtaining interest from individual users. This allows people to freely express themselves and to influence others on the network. Therefore, the digital channels have become extremely important, and it is a source of an extremely large and important amount of information.

The information that the users leave on the digital channels related to daily lives, and it is necessary to analyze the obtained opinion in order to automatically manage the public review and support decision making. For instance, the comments collected from Twitter have been utilized to predict the election outcomes [1].

In general, the information collected from the digital channels helps the agencies or organizations to gain insight: (1) The customers use this information for searching or referring purposes before they make a decision about using a product or service; (2) The service providers can also use these sources of information to evaluate their products, thereby they are able to make improvements more relevant to users, bring higher profits, and avoid unfortunate risks. Especially, as the businesses have a new product, they want to launch it in the market. It is extremely necessary to get the customer feedbacks or surveys; (3) The authorities can use this information to find out the comments and attitudes of the community in order that they can promptly adjust and announce policies to make them more reasonable.

To solve the mentioned matters effectively and quickly, we need to take advantage of artificial intelligent in analyzing the collected information such as comments, reviews, and opinions that the customer or users leave over the digital channels. We consider this issue as a text classification which aims to assign the text to various categories. The application of text classification comprises topic classification [2], question classification [3], and sentiment analysis [4]. There are various techniques proposed to solve these problems so far. Especially, the deep learning technique has become popular in such text classification as CNN [5], recurrent neural networks (RNN) [6], and other complicated approaches. With the deep learning-based approach, the model takes a text as an input to generate text representation at the output. This output is then passed to the softmax or sigmoid layer to produce the probability of each label.

For sentiment analysis, the main objective is to classify the text content to positive, neutral, or negative sentiment. There are numerous surveys and researches published

in this area, and the sentiment classification has been categorized into three levels like aspect-level, sentence-level, and document-level [7, 8]. The levels of sentence and document are interesting and challenging. However, the aspect-level is more complex because it performs finer-grained insights. The sentiment analysis can be developed relied upon various approaches like approaches based on lexicon, machine leaning methods, and hybrid methods [9–11].

Besides, we have found many published researches for analyzing Vietnamese texts [12–15]. Most existing approaches for Vietnamese texts are developed by using BERT, PhoBERT [16], CNN, CNN + long short-term memory (LSTM), and ensemble learning. The pre-trained BERT utilizes the technique of word separation for English, so it is unaware of the difference between Vietnamese syllables and word tokens. Meanwhile, the PhoBERT is a novel model for Vietnamese language. This is the reason we choose the PhoBERT-base in combination with CNN to develop a sentiment analysis for Vietnamese texts, and the analysis is conducted based on the sentence level. With this combination, the proposed model consists of two-fold: The first one is that the model can effectively obtain local information in the texts by embedding the last four hidden layers of PhoBERT into several filters and convolutional layers of the CNN; the other is that the local representation obtained from PhoBERT is fed into transformer encoder. This helps model to be capable of using the self-attention technique to generate the final representation more focusing on the text segments.

The paper is structured as follows. The related works are presented in the next section. The proposed method is presented in Sect. 3. Section 4 presents the experimental results. Conclusion and future improvement are discussed in Sect. 5.

2 Related Works

We briefly review the existing sentiment analysis approaches for English and Vietnamese texts which are closely relevant to our work.

2.1 Sentiment Analysis for English Text

Deep neural network has recently achieved good performance in the field of natural language processing (NLP). CNN-LSTM presented in [17] uses a combined kernel from multiple branches of CNN along with LSTM layers to analyze movie reviews. By minimizing model overfitting, the authors show that this approach give high performance. Yu et al. [18] have proposed a sentiment analysis for public news in which the authors refine the key news from long and large news data. The k-means clustering is used to divide the high-dimensional vector into k categories. Another approach based on CNN has been proposed by Kim and Jeong [19]. The proposed network consists of two convolutional layers, a pooling layer, an embedding layer,

and a fully connected layer. The authors demonstrate that their model is better than the other deep learning-based ones. Suciati and Budi [20] present a comparison of deep learning machine and learning in sentiment analysis, and this comparison is also conducted in emotion classification with multilabel classification. For machine learning, the techniques like binary relevance, classifier chain, label powerset and random forest, decision tree, support vector machine, and extra tree classifier are used. Regarding deep learning, the authors utilize such techniques as gated recurrent unit, bidirectional long short-term memory.

Ghorbani et al. [21] have proposed a method by integrating CNN and LSTM with the aim of identifying word polarity on the Google cloud platform. The authors also make use of word embedding technique which learns features from the CNN layers. These obtained features are directly fed into a bidirectional LSTM layer to gain the dependencies of long-term feature. The proposed method gives high accuracy. Another deep learning approach presented by Peng and Kexin Zhang [22] introduces three deep learning networks for sentiment analysis. These networks comprise RNN, LSTM, and CNN, and the public IMDB movie reviews are used to train and assess the models. An aspect-based sentiment analysis [23] is developed by using the pre-train BERT which produces contextual word representations. The author also uses only one sentence pair-based classifier from BERT to deal with both sentiment analysis and aspect analysis. FastText and BERT have been proposed by Kastrati et al. [24] in which the authors show that BiLSTM used together with an attention mechanism give highest performance on the sentiment analysis for Facebook opinion. The work presented in [25] shows the comparison between various deep learning networks and LSTM ones. The authors have performed several combination of the networks and concluded that the usage of BERT achieves the best results. However, the time consuming is high for training process. Recently, a combination of BERT and CNN has been proposed by Dong et al. [26]. In this work, the authors take advantage of pre-trained BERT to extract global features of the text vectors. The features obtained from the BERT layer are then fed into the CNN network for sentiment classification.

2.2 Sentiment Analysis for Vietnamese Text

We have observed that the sentiment classification for Vietnamese text is less studied than that of English text. Nguyen et al. [27] present a work in which the authors make use of several techniques like maximum entropy, Naive Bayes, bidirectional long short-term memory, long short-term memory. This work shows that the bidirectional long short-term memory model outperforms the other ones. Another work presented in [28] is based on a combination of traditional and deep learning approaches. The authors show that the proposed method achieves high accuracy than using each one alone. The sentiment analysis for student feedback [29] has been proposed by utilizing LSTM and dependency tree-LSTM. The authors present that the final hidden state vectors of LSTM and dependency tree-LSTM models combined with a support vector machine give high performance than the LSTM model.

Khai Tran and Thi Phan [30] proposed a sentiment analysis by using the ensemble learning model. This kind of model consists of various components like language features, sentiment shifting, and statistical techniques. The word embedding is also adapted to the deep learning model for solving sentiment factors in text. The proposed model is better than the others which are based on the machine learning and deep learning techniques. Another approach proposed to address the feature extraction at the sentence level has been presented by Pham and Kieu [31]. The gate framework is used for building a rule-based system. The student feedback corpus has been presented by Nguyen et al. [32] in which the authors have built a Vietnamese students' feedback corpus for the purpose of sentiment analysis. Besides, the authors also take advantage of two existing tools like Datumbox framework and Stanford classifier to perform the classification on their corpus.

3 Proposed Method

To extract the sentiments like positive, negative, or neutral from a given opinion, we need several steps and solve various natural language processing matters. In general, after collecting and extracting the feedback from several sources in diverse formats, we need to convert it to normal text, and this text is processed by using natural language processing (NLP) techniques. Below is the main steps of our work.

- Text sentence preprocessing.
- The text sentence is segmented into word level (in a form of Vietnamese text) before putting into PhoBERT.
- Encoding word by using PhoBERT tokenizer.
- Putting the encoding text sentence into the model with attention mask.
- Taking the output vectors at the last four layers of PhoBERT and concatenating them as a sentence feature.
- The sentence feature is then fed into the CNN model for text classification.

3.1 Data Preprocessing

The text obtained from the customer feedback needs to be processed in order to maximize the extracted features which is better for further steps and to obtain a clean text. The raw feedback content often contain noise and has not been segmented into a normal form of Vietnamese text. Unlike English text, a Vietnamese word can be made up of more than one sound (about 85% of Vietnamese word types are made up of at least two syllables). For example, the word "giáo_trình" (corresponding to "syllabus" in English) is made up of two sounds, namely "giáo" and "trình". Meanwhile, the two single words like "giáo" and "trình" can have different meanings. Therefore, separating Vietnamese word is an important step that we need to take before putting

Table 1 List of emoticons

Emoticons	Meaning	Regex	Replacement
:), :), :-), (:, (:, (-:, :')	Smile	(:\s?\)|:-\)|\(\s?:|\(-:|:\'\))	pos_val
;-), ;), ;-D, ;D, (;, (-;	Wink	(:\s?\(|:-\(|\)\s?:|\)-:)	pos_val
:-(, : (, :(,):,)-:	Sad	(:\s?\(|:-\(|\)\s?:|\)-:)	neg_val
:,(, :'(, :"(	Cry	(:,\(|:\'\(|:"\()	neg_val

data into the next step. In this work, we utilize the package named VnCoreNLP [33] to segment the Vietnamese words.

- The raw text (syllable level): "nhiệt tình giảng dạy, gần gũi với sinh viên"
- The segmented text (word level):
  "nhiệt_tình giảng_dạy, gần_gũi với sinh_viên"

We then apply a number preprocessing steps to minimize the redundant information from the dataset and to reduce its size. These steps are as follows.

- Replace two or more dots with space.
- Substitute two or more spaces with a single space.
- Convert the text to lower case.
- Remove spaces from the end of texts.
- Replace all emoticons with the value of "pos_val" or "neg_val", which is given in Table 1.
- Convert two or more letter repetitions to two letters.

3.2 PhoBERT-Based CNN Model

The pre-trained PhoBERT model is the well-known language models for Vietnamese text. This model can be fine-tuned, or used directly as a feature extraction for various textual tasks. The proposed model comprises two main processes. The first one is the PhoBERT-based model in which the text is passed through several self-attention layers. This process aims to obtain the contextualized vector representation of the text. With these self-attention layers, the output value corresponding to each input token is acquired. The other is the CNN model which is used to extract the local information (features) from the output of the PhoBERT's layers. We employ the transformer encoder to integrate the local information obtained by CNN model and the representation of the entire text obtained by PhoBERT. By using the self-attention mechanism, the model is capable of making the final sentiment representation concentrating on the important information of the sentiment. The architecture of CNN, which is depending on PhoBERT, is depicted in Fig. 1.

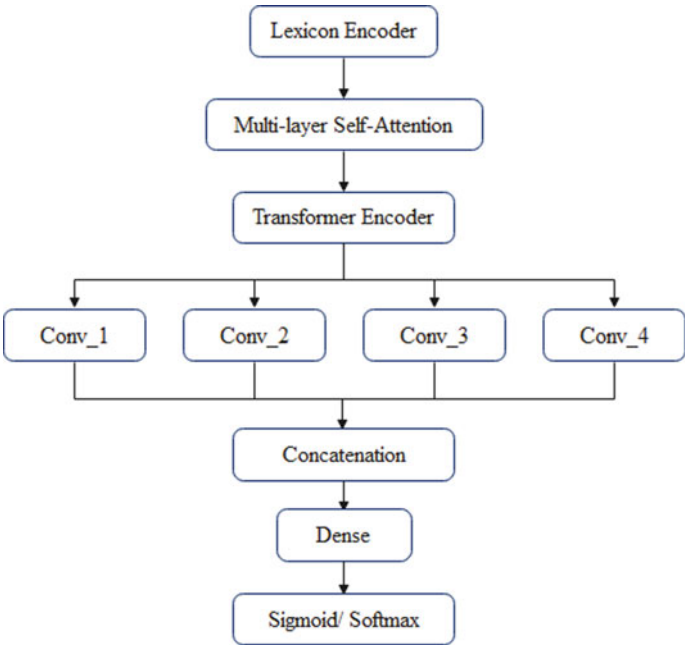


Fig. 1 The workflow of PhoBERT-based CNN

For encoder, we utilize the encoder with multiple layers to map the input to a series of contextualized embedding vectors $V = \{c, \{T_1, T_2, \ldots, T_n\}, s\}$ where c is the contextual representations for [CLS], s is the contextual representations for [SEP], and T_i is the contextual representation corresponding to the ith token. We then use the transformer encoder to map the obtained representations to the representation of entire text. The PhoBERT input is a text, and the last four hidden layers of the PhoBERT is then combined to obtain the representations of size ($768 \times 4 \times 100$), where 100 is the maximum length of each text sample. These vector representations are then parallelly passed to the convolutional layers (1D) of four different sizes like (768×2), (768×3), (768×4), and (768×5). The number of filters in each size is 256. The output of the last four layers of PhoBERT is passed into each kernel. Each kernel then has four channels, and the convolutional operation is then applied on these kernels. The output of this step is fed into the ReLU activation function and max pooling operation. Finally, the output of the max pooling operation is concatenated and flattened before passing into the last dense layer. The sigmoid function is used at the last dense layer if we want to get a binary classification. Otherwise, in case of multi-classification, the activation function is set to softmax.

The PhoBERT-based CNN classifier is defined as follows.

$$\hat{y} = \text{softmax}(\text{ReLU}(Wc + b)) \tag{1}$$

where W is the weights and c is the representations obtained from the transformer encoder and CNN layers.

The cross-entropy is used as a loss function in this work.

$$L = \sum_{i} y_i \log \hat{y}_i \tag{2}$$

where y_i is the sentiment class and $\hat{y}_i$ is the prediction of the ith corresponding sentiment.

4 Experimental Results

4.1 Dataset

We employ three public datasets to evaluate the performance of our model. These datasets consists of sentiment analysis. The information of these datasets is depicted in Table 2.

- Student feedback [32]: This dataset contains the student feedback in which the students give opinions about a variety of issues. For sentiment-based task, the training set comprises 11.427 comments with three classes like positive, negative, and neutral. The testing set consists of 3.167 comments.
- Food reviews [34]: This is a dataset of food reviews which have been labeled with two types of positive and negative sentiment. The comments are taken from the Foody site. The classification of positive and negative label is based on the average score of each comment.
- Ntc_sv [35]: It contains about 50,000 samples which is obtained from the reviews of food and restaurants on Foody site. The reviews are labeled with negative and positive sentiment.

We use the datasets with different comments to better evaluate the various models like traditional (SVM and XGBoost) and deep learning (CNN, BERT, and PhoBERT-based CNN) ones. The comments made in the datasets are separated with their

Table 2 The information of the three datasets for sentiment analysis

Datasets	Type	Train size	Validation size	Test size	Labels	Length (avg.)
Student feedback	Sentiment	11,427	–	3167	3	13.35
Food review	Sentiment	30,000	10,000	10,000	2	94.77
Ntcsv	Sentiment	40,761	–	10,000	2	78.25

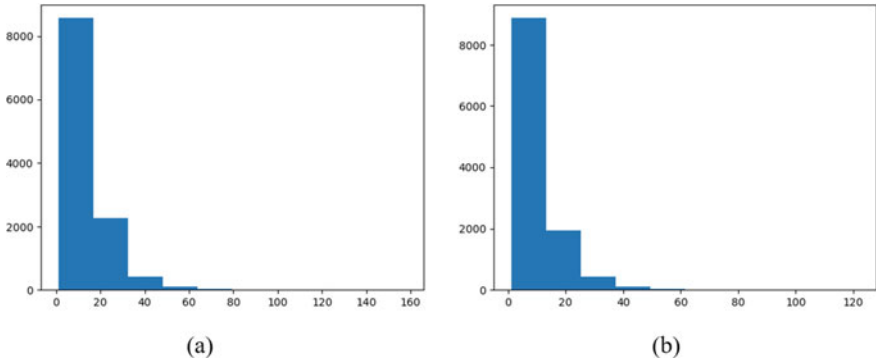


Fig. 2 Number of words of each comment before (**a**) and after (**b**) transforming into Vietnamese normal form

corresponding sentiments (labels) by a comma. The training datasets are in a csv file. For the datasets with two classes, the sentiment is either 0 (negative) or 1 (positive). The sentiment is assigned a value of 0 (negative), 1 (neutral), or 2 (positive) in case of three classes. Similarly, the test sets are also in a csv file. In case of the datasets without providing the validation set, the validation set is then made by splitting 20% samples from the training set.

4.2 Experimental Setup

The proposed network is implemented by using TensorFlow implementation of PhoBERT. We choose the pre-trained model of $\text{PhoBERT}_{\text{base}}$ (12-layer, 768-hidden, 12-heads, and 135M parameters) because of limited computational performance. The size of a batch is set to 32, and the maximum sequence length of PhoBERT-based CNN is set to 100 for all datasets in our experiments. The dropout operation is added after the transformer encoder with multiple layers, the transformer encoder, and CNN layer to minimize the overfitting. We optimize the model by using Adam with the learning rate of $2e-5$. The number of epochs is assigned to 10.

4.3 Results

Figure 2 shows that the number of words of each sentence can be reduced after transforming to the normal form of Vietnamese text. The statistics are conducted on the student feedback dataset. The process of removing noises from the comments also leads to a reduction in length.

Fig. 3 Frequency of negative and positive words appeared in the training set of the student feedback dataset

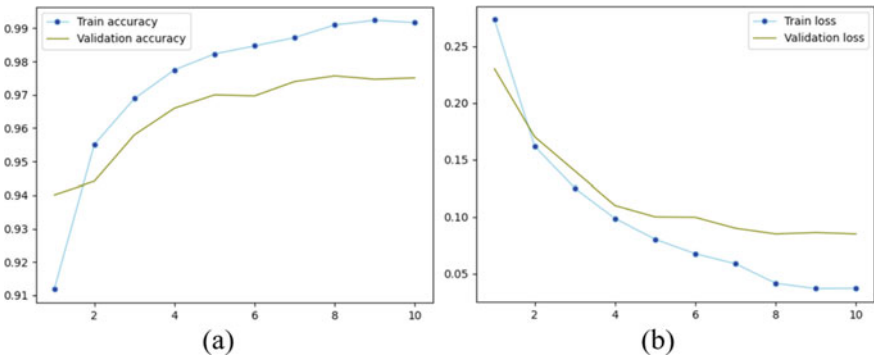


Fig. 4 Accuracy (**a**) and loss (**b**) corresponding to train and validation set

The frequency of words appeared in negative and positive comments is depicted in Fig. 3a and b, respectively. As we can observe that the feedbacks classified as negative frequency contain words like “nên”, “không”, “thực_hành”, “khó_hiểu” . However, the less frequency words seem to offer more insight like “giảng_dạy”, “chất_lượng”, “quá_nhiều”. Meanwhile, the positive feedbacks contain common words like “ “nhiệt_tình” ”, “tận_tình”, “giúp sinh _viên”. They also mention works like “quan_tâm đến”, “có rất”, “truyền_đạt tốt”, which are highly valued by the students.

The neutral comments are generally descriptive. Even though these comments do not express a particular sentiment, they make clear on topics that are related to the students.

Table 3 Results of sentiment analysis obtained from the student feedback dataset

Approach	Precision	Recall	F_1-score
SVM	0.9058	0.8014	0.8504
XGBoost	0.8714	0.8136	0.8415
CNN	0.8931	0.8705	0.8817
CNN+LSTM	0.8517	0.8903	0.8705
BERT-base	0.9414	0.8502	0.8935
PhoBERT-based CNN	**0.9602**	0.8725	**0.9143**

Table 4 Results of sentiment analysis obtained from the Ntc_SV dataset

Approach	Precision	Recall	F_1-score
SVM	0.8851	0.9031	0.8940
XGBoost	0.8810	0.9029	0.8918
CNN	0.8722	0.8526	0.8623
CNN+LSTM	0.8913	0.8638	0.8773
BERT-base	0.8814	0.8224	0.8509
PhoBERT-based CNN	**0.9375**	0.9024	**0.9196**

The accuracy and loss of training process on the student feedback dataset are depicted in Fig. 4. The curves of train and validation accuracy show that the proposed model tends to be overfit after reaching the highest accuracy at the 8th epoch. However, the evaluation accuracy meet our expectation for this work, and the classification conducted on the testing sets give better performance than other approaches. The experimental results of various methods are presented in Table 3. The figures show that the deep learning techniques outperform the conventional ones. In addition, during the model training, we have observed that the training time of PhoBERT-based CNN is much longer than that of XGBoost, SVM, and even CNN method.

The experimental results conducted on the testing set of the Ntc_sv are illustrated in Table 4. The BERT-based method also gives better result in sentiment classification than other approaches, especially the traditional approaches like SVM and XGBoost. However, the PhoBERT-based CNN gives the best result. For the food review dataset, due to the computational time, we only present the results of our method, without comparing with other approaches. The precision, recall, and f_1-score are 0.9405, 0.8823, 0.9104, respectively.

Compared with BERT-base and other methods, the PhoBERT-based CNN performs better than on three datasets. This illustrates that the CNN model is effectively used to improve the features obtained from the last four layers of PhoBERT. By using CNN, the model is capable of obtaining new text representation from the feature vectors extracted by PhoBERT. This makes the model to concentrate on the local information in the sentiment. The self-attention mechanism is also another advantage of PhoBERT which helps the sentiment representation concentrate on the important

Table 5 Comparison of BERT-based and PhoBERT-based approach

Approach	Precision	Recall	F_1-score
Conducted on the student feedback dataset			
UIT-VSFC [32]	0.8771	0.8866	0.8794
Our approach	**0.9602**	0.8725	**0.9143**
Conducted on the Ntc_sv dataset			
BERT-base [14]	0.8813	0.9402	0.9090
BERT-TextCNN [14]	0.8885	0.9314	0.9094
Our approach	**0.9375**	0.9024	**0.9196**

information of the sentiment. Focusing on the local information and text segments makes up the improvement in our approach. Although the training time is longer than the CNN model, the proposed approach obtains better results, without taking too much computational cost.

Table 5 shows the comparison results with the existing works in which the testing is conducted on two public datasets. The figures illustrate that the PhoBERT-based CNN outperforms the state-of-the-art approaches.

5 Conclusion

We have proposed a new model for sentiment classification by combining the PhoBERT and CNN model. The PhoBERT model is used to extract the feature vector which contains full information of bidirectional context. This information is obtained through the self-attention mechanism in each transformer block. The feature vector obtained from the PhoBERT model is then fed into the convolutional layers for extracting the local information in the text. Combining the representation of the whole text and local representation of the text gives good results in classifying the Vietnamese sentiment. However, the model needs to be further improved in order that the preprocessing stage is capable of replacing meaningless words in a sentence. The new text representation from the feature vectors extracted by PhoBERT is also another point which is considered in the future works.

References

1. O'Connor B, Balasubramanyan R, Routledge B, Smith N (2010) From tweets to polls: linking text sentiment to public opinion time series
2. Wang S, Manning CD (2012) Baselines and bigrams: simple, good sentiment and topic classification. In: Proceedings of the 50th annual meeting of the association for computational linguistics: short papers, vol 2

3. Zhang D, Lee WS (2003) Question classification using support vector machines. In: Proceedings of the 26th annual international ACM SIGIR conference on research and development in information retrieval
4. Maas AL, Daly RE, Pham PT, Huang D, Ng AY, Potts C (2011) Learning word vectors for sentiment analysis. In: Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies, vol 1, pp 142–150
5. Zhang X, Zhao J, Lecun Y (2015) Character-level convolutional networks for text classification
6. Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. In: NIPS 2014 workshop on deep learning, Dec 2014
7. Behdenna S, Barigou F, Belalem G (2016) Sentiment analysis at document level. In: Smart trends in information technology and computer communications
8. Ha H, Prasad P, Maag A, Alsadoon A (2019) Deep learning for aspect-based sentiment analysis: a comparative review. Expert Syst Appl
9. Madhoushi Z, Hamdan AR, Zainudin S (2015) Sentiment analysis techniques in recent works. In: 2015 Science and information conference (SAI)
10. Sankar H, Subramaniyaswamy V (2017) Investigating sentiment analysis using machine learning approach. In: 2017 International conference on intelligent sustainable systems (ICISS)
11. Thakkar H, Patel D (2015) Approaches for sentiment analysis on twitter: a state-of-art study
12. Kieu BT, Pham SB (2010) Sentiment analysis for Vietnamese. In: 2010 Second international conference on knowledge and systems engineering
13. Nguyen K, Nguyen V, Nguyen P, Truong T, Nguyen N (2018) UIT-VSFC: Vietnamese students' feedback corpus for sentiment analysis
14. Nguyen QT, Nguyen TL, Luong NH, Ngo QH (2020) Fine-tuning BERT for sentiment analysis of Vietnamese reviews. In: 2020 7th NAFOSTED conference on information and computer science (NICS)
15. Nguyen-Thanh T, Tran GT (2019) Vietnamese sentiment analysis for hotel review based on overfitting training and ensemble learning. In: Proceedings of the tenth international symposium on information and communication technology
16. Nguyen DQ, Tuan Nguyen A (2020) PhoBERT: pre-trained language models for Vietnamese. In: Findings of the association for computational linguistics: EMNLP 2020
17. Yenter A, Verma A (2017) Deep CNN-LSTM with combined kernels from multiple branches for IMDB review sentiment analysis. In: 2017 IEEE 8th Annual ubiquitous computing, electronics and mobile communication conference (UEMCON)
18. Yu X, Zhong C, Li D, Xu W (2020) Sentiment analysis for news and social media in covid-19. In: Proceedings of the 6th ACM SIGSPATIAL international workshop on emergency management using GIS
19. Kim H, Jeong YS (2019) Sentiment classification using convolutional neural networks. Appl Sci
20. Suciati A, Budi I (2020) Aspect-based sentiment analysis and emotion detection for code-mixed review. Int J Adv Comput Sci Appl
21. Ghorbani M, Bahaghighat M, Xin Q, Özen F (2020) ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing. J Cloud Comput Adv Syst Appl
22. Peng C, Kexin Zhang DZ (2020) Sentiment analysis using deep learning approach. J Artif Intell
23. Hoang M, Bihorac OA, Rouces J (2019) Aspect-based sentiment analysis using BERT. In: Proceedings of the 22nd Nordic conference on computational linguistics
24. Kastrati Z, Ahmedi L, Kurti A, Kadriu F, Murtezaj D, Gashi F (2021) A deep learning sentiment analyser for social media comments in low-resource languages. Electronics
25. Colón-Ruiz C, Segura-Bedmar I (2020) Comparing deep learning architectures for sentiment analysis on drug reviews. J Biomed Inform
26. Dong J, He F, Guo Y, Zhang H (2020) A commodity review sentiment analysis based on BERT-CNN model. In: 2020 5th International conference on computer and communication systems (ICCCS)

27. Nguyen P, Hong T, Nguyen K, Nguyen N (2018) Deep learning versus traditional classifiers on Vietnamese students' feedback corpus. In: 2018 5th NAFOSTED Conference on information and computer science (NICS)
28. Nguyen HQ, Nguyen QU (2018) An ensemble of shallow and deep learning algorithms for Vietnamese sentiment analysis. In: 2018 5th NAFOSTED conference on information and computer science (NICS)
29. Nguyen VD, Nguyen KV, Nguyen NLT (2018) Variants of long short-term memory for sentiment analysis on Vietnamese students' feedback corpus. In: 2018 10th International conference on knowledge and systems engineering (KSE)
30. Khai Tran T, Thi Phan T (2019) Deep learning application to ensemble learning-the simple, but effective, approach to sentiment classifying. Appl Sci
31. Pham S, Kieu B (2010) Sentiment analysis for Vietnamese. In: International conference on knowledge and systems engineering
32. Nguyen KV, Nguyen VD, Nguyen PXV, Truong TTH, Nguyen NLT (2018) UIT-VSFC: Vietnamese students' feedback corpus for sentiment analysis. In: 2018 10th International conference on knowledge and systems engineering (KSE)
33. Vu T, Nguyen DQ, Nguyen DQ, Dras M, Johnson M (2018) VnCoreNLP: a Vietnamese natural language processing toolkit. In: Proceedings of the 2018 conference of the North American chapter of the association for computational linguistics: demonstrations
34. https://streetcodevn.com/blog/dataset
35. https://www.aivivn.com/contests/1

Evaluation of Cloud Databases as a Service for Industrial IoT Data

Theodosios Gkamas, Vasileios Karaiskos, and Sotirios Kontogiannis

Abstract Applications using IoT sensory data, such as in Industry 4.0, are a classic example of an organized database. This paper focuses on evaluating three types of DBMS, MongoDB, PostgreSQL using JSON and the relational PostgreSQL, measuring average, jitter, and loss of response Time and achieved throughput. Three scenarios were thoroughly tested, (i) data insertions, (ii) select/find queries, and (iii) queries related to correlation functions. Experimentations concluded that MongoDB is between 19–30% faster than Postgres in the insert queries, achieving 51–55% higher throughput. Additionally, relational Postgres is x4 times faster than MongoDB and x2 times faster than Postgres JSON in the selection queries, achieving 31–35% higher throughput. Finally, the two versions of Postgres performed equally concerning response time in the correlation function queries, while both of them outperformed MongoDB by x3.6 times. Contrariwise, in the correlation function queries, MongoDB achieved 19–24% higher throughput than both versions of Postgres.

Keywords Database systems performance evaluation · MongoDB · PostgreSQL · Industrial systems · IoT data · Industrial IoT

1 Introduction

Industry 4.0 [1] is nowadays a reality that accelerates and evolves at a fast pace. Information has no more a single source, but on the contrary comes from several machines, devices, sensors, or embedded systems [2]. With the appearance of the Industrial Internet of Things (IIoT), such as wireless sensors and actuators, we have passed from the digital era to the intelligent one [3]. For the intelligent industry,

T. Gkamas · V. Karaiskos · S. Kontogiannis (✉)
Laboratory of Distributed Micro-computer Systems, Department of Mathematics, University of Ioannina, 45500 Ioannina, Greece
e-mail: skontog@uoi.gr

T. Gkamas
URL: http://kalipso.math.uoi.gr/microlab/en/

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_25

advanced manufacturing based on network and application data must be considered. Those smart objects can communicate with each other [4] and can provide us with tools such as big data analytics and cloud computing [5], leading us to the creation of Cyber-Physical Production Systems (CPPSs) [6].

Appropriate control systems with Internet connection are required to handle the measurements of a set of IIoT sensors. Two main concepts exist in the industry, DCS and PLCs. A decentralized control system (DCS) is a control method with some independent CPUs. Each CPU controls individual parts of the process, so if one fails, the other CPUs will continue executing their function without disturbance. DCS is more suitable for continuous processes with lots of analog signals and complicated proportional–integral–derivative (PID) control loops, such as processes in an industrial environment or oil refinery or power plant. Programmable logic controller (PLC) does not imply the approach of centralization or non-centralization on its own. Practically, a PLC is mainly used for centralized control, meaning that a single PLC or CPU will control the whole process. PLCs have become more progressive and mature these days and can even control the PID loops and analog signals (by using extensions for analog input). Even with multiple PLCs in an industrial environment, we still need to create a managerial and control system. DCS database would be available for the creation of the monitoring and system. In contrast, individual PLC databases would need to be created in the SCADA system software in the PLC systems.

Distributed database systems have been proven compelling in the effort to deal with this data cataclysm. These storage systems are prominent by two key factors: (a) system scalability and (b) interactive performance: very fast response times to client requests. According to [7], there are challenges in managing and querying the massive scale of spatial data. Benchmarks play a crucial role in examining the performance and functionality of spatial databases both for commercial users and developers.

MongoDB is an open-source document-based NoSQL database which is supported commercially since 2007 [8]. Although MongoDB is non-relational, it implements many features of relational databases, such as sorting, secondary indexing, range queries, and nested document querying. Operators like insert, read, update and remove, and manual indexing, indexing on embedded documents, and index location-based data are also supported. MongoDB documents are serialized consistently as JavaScript Object Notation (JSON) objects and stored internally using a binary encoding of JSON called BSON [9].

On the other hand, PostgreSQL is an open-source object-relational database system, having more than 35 years of effective development on the core platform [10]. PostgreSQL has gained in popularity for its proven architecture, reliability, data integrity, robust feature set, and extensibility [10]. It supports a wide variety of data types, including JSON or JSONB types. Primarily, JSON is used in one of the benchmarks for using OLAP queries reflecting MongoDB. For datasets stored in memory, it is possible to access in-memory columnar data stores in Postgres through Cstore [11] (another plug-in extension) to speed up aggregations in memory. Another useful

extension is pg-stat-statements [12] which tracks statistics on the queries executed by a Postgres database.

In this paper, we compare the performance in terms of response time, achieved throughput and loss between MongoDB [13], and two variations of the open-source ORDBMS, PostgreSQL [14], the relational one, and the one using JSON files. We did not include Postgres JSONB files since we do not perform records updates in our use cases. The rest of this manuscript is organized as follows. Section 2 contains the bibliography about industry 4.0 and relational and non-relational databases. Section 3 presents the Industry 4.0 architecture and capabilities. Section 4 describes the selected performance evaluation using industrial data and the obtained results. Finally, Sect. 5 concludes this paper.

2 Related Work

Chen et al. [3], proposed a hierarchical architecture of the smart factory and additionally the key technologies used for the physical assets, the network, and data application layers. Padovano et al. [2] designed an application prototype focused on a digital twin as a service provider for Industry 4.0 in order to add perception and handling on the physical system.

According to [15], the Postgres database was measured to be 4–15 times faster than MongoDB in transaction performance testing conducted by OnGres [16]. Moreover, Postgres prevails in JSON-based online analytical processing (OLAP) tests designed to operate on document-based data. Furthermore, tests on online transaction processing (OLTP) workloads, using the standard Sysbench [17] benchmark, have shown that Postgres is 3 times faster on average than MongoDB in running multi-document, surveying both performance and latency. The tests concluded that if someone wants a fast performance with lower latency, then Postgres is the leading choice. Concerning cybersecurity PostgreSQL has a full range of opportunities and works with all major cloud providers although operational and developers tools differ according to the cloud provider.

If using a distributed database containing data of any structure that these data shall scale out in terms of volume of traffic or size of the data, then MongoDB is suitable for that task. Moreover, if you want a multi-cloud database, which operates the same on every public cloud, MongoDB shall be selected. MongoDB stores data as documents in a BSON representation without needing to change the structure of the documents, as is the case with relational databases. Query performance in MongoDB can be increased by creating indexes on (sub)documents fields. Taking advantage of its distributed form of the database, MongoDB allows the user to scale out, providing us the possibility to create a large number (i.e., hundreds) of smaller machines. Furthermore, a modern cybersecurity suite is available in MongoDB either for on-premises or cloud versions with the same configuration for the three major cloud providers.

Makris et al. [7] implemented a PostgreSQL including the PostGIS extension system [18] with a 5-node cluster setup, versus single node implementation with the use of indexes. Experimentations concluded that PostgreSQL is 4 times faster in response time in most cases and presents an average speedup around 2 times in the first query, 4 times in the second query, and 4.2 times in the third query for a 5-node cluster. Additionally, the authors observed that the average response time is significantly reduced at half with indexes almost in all cases, while the reduction is significantly lower in PostgreSQL. Klein et al. [19] presented a study among three NoSQL databases for a large, distributed healthcare organization. While the algorithm considered consistency, availability, partition tolerance (CAP) trade-offs, and other quality attributes that influence the selection decision, they reported on the performance evaluation method and results.

To summarize, Postgres is the best option if we have a relational data model that will not change with time. On the contrary, if built-in scalability is the main IIoT design goal, MongoDB can firstly scale horizontally with native sharding, secondly can scale out vertically by adding faster nodes and memory, and thirdly can offer fail-over and replication mechanisms. Unfortunately, PostgreSQL supports replication clustering and load balancing in terms of connections clustering and not automated table sharding as MongoDB since it highly affects the database availability. PostgreSQL sharding is supported on top of partitioning by allowing any given partition of a table to be hosted by a remote server.

3 IIoT Architecture Based on Industry 4.0 Standards

Installing sensors in an industrial environment that communicate with the corresponding concentrators and a PLC or DCS controller allows transferring data from the controller to the cloud system, with the use of NarrowBand IoT (NB-IoT) or 3G/4G protocols, to the MQTT broker which continuously sends measurements, as encrypted (Base 64) encoded JSON strings, indirectly to the database service via a bidirectional connection with a https server. Primarily to data storing, an intermediate service MQTT-to-DB exists, which decodes the obtained encrypted IIoT data and stores them to the database cluster.

Afterward, the server offers a variety of applications to users, through a push pop http API, allowing to process, analyze, and manage the collected sensory information via a Web interface or a mobile application with the aid of Grafana application for visualization purposes and the derived statistics. Such an application is installed in any personal device, such as a PC, tablet, or smartphone. The application can respond to incidents, apply AI smart agents, and bots producing mass notifications and other capabilities, such as enhanced AR applications, monitoring near real-time sensory measurements, and historical data.

4 Database Performance Evaluation

4.1 Dataset Overview

The dataset used to perform the DBMS testing scenarios will be an industrial compressor non-indexed sensory data collection of 400 K records split in tables (for PostgreSQL) or collections (for MongoDB) of equal size. The number of tables or collections is calculated according to the 1% of the number of queries per second (QPS) executed on each DBMS. The dataset will consist of randomly generated data utilizing the same procedural method. Each record consists of a unique identifier that each DBMS would generate at successful data insertion. A "tagid" field containing the tag identifier for the simulated device, a sub-collection named "rxInfo" containing the simulated device's informational fields, and a sub-collection named "data" containing the sensor measurements. The informational fields will consist of the received signal strength indicator (RSSI), the GPS location coordinates of the physical device, the collector's gateway ID, and the timestamp of each measurement. The sensor measurements will consist of the rotation per minute (RPM) measurements, temperature measurements around the engine axis and engine frame, measured in °C, and 3D Cartesian plane vibration vectors of four acceleration sensors placed on the NW, NE, SW, SE of the simulated engine frame, measured in m/s^2.

The constructed dataset has been used to consider real-world applications and needs, utilizing a variety of information necessary for daily operations in an industrial setting. The total size of the records was decided to be 400 K in order to fit both the criteria of being large enough for an IoT database set consisting of millions of network packet transactions between the IoT devices and database servers and small enough to perform a simulation of a real-world deployed platform, as a part of the proof of concept.

4.2 Evaluation Metrics

Following the example of [20], the authors used the following metrics: (a) Average and (b) Jitter of response time, (c) Average and (d) Jitter of achieved throughput and (e) the Loss, meaning the database drops.

4.3 Performance Evaluation and Experimentation Using IIoT Data

The selected clustering system consists of 20 client nodes and a server. Each client node is a PC running Linux Slax 4.9.0-11-686 SMP Debian 4.9.189-3 in a virtual box with AMD Ryzen 32200G CPU and 1.5 GB RAM, while the server is a PC

running Linux Ubuntu 20.04.3 with Intel Xeon E5-2640 2.5GHz (24 cores) CPU and 32GB RAM. The DBMS testing scenarios were designed to simulate real-world case scenarios. For this experiment, three separate use cases were selected:

- Non-relational MongoDB document-styled records.
- Non-relational JSON PostgreSQL document-styled records.
- Relational PostgreSQL table row-styled records.

For each case, three separate scenarios were selected for testing purposes:

(Q1) Insert query: Data records are generated and inserted into the DBMS using the network of 20 computers in multi-threaded order consisting of four test cases of $T = \{100, 500, 1\,\text{K}, 10\,\text{K}\}$ QPS on each of the DBMS use cases.

(Q2) Select query: Utilizing the databases constructed in the previous test scenarios, selection queries will be performed on the DBMS to fetch data that match the filtering condition. For this case, the query's dataset returns all records (lines/documents) from a single table/collection. All fields are procured only if the "tagid" is matched from the variables set that the executing algorithm produces. The variable set is randomly selected from a list containing all tables/collections paired with each unique "tagid" within said table/collection. The queries will be performed in a similar structured algorithm as in the Q1 case, utilizing the same network in similar multi-threaded order with for test cases of T QPS up to the maximum network throughput.

(Q3) Correlation function query: In order to perform a complex type of querying, the same setup as in the Q2 case is used, differentiating in the expected dataset containing one single record, only the arithmetic mean of each "data" sub-collection field (sensor measurements). Data correlation queries will be performed in the same network as before, in similar multi-threaded order with four testing cases of T QPS up to the maximum network throughput.

All testing scenarios will be evaluated and confirmed by replicating the same results. Results are stored in JSON trace files for ease of access and processing.

4.4 IIoT Data Experimentation and Experimental Results

Initially, to the experimentation, we shall introduce some generic comments. The size of each of the three databases is a total of 400 K records. Every scenario is executed 100 K times for verification.

Scenario (Q1)—IIoT Data Insertion: MongoDB is 19% faster than PostgreSQL using JSON and 30% faster than relational Postgres in the insert queries, achieving 51 and 55% higher throughput than PostgreSQL JSON and relational PostgreSQL, respectively. At the same time, the response time jitter is around 0.038 s, and the throughput jitter is around 45–52 QPS.

Scenario (Q2)—IIoT Data Selection: The authors fetch 1% of 400 K records in each database (around 4 K records). The relational PostgreSQL is four times faster than MongoDB and two times faster than PostgreSQL using JSON in the selection

Table 1 Average, jitter and loss of response time and throughput in MongoDB, Postgres JSON and relational postgres for the representative case of 10 K client throughput

Scenarios	Database type	Average response time (s)	Jitter of response time	Response loss (%)	Average achieved TP	Jitter of achieved TP	TP loss (%)
Q1	**MongoDB**	**0.121**	**0.038**	**0.00**	**476**	52	**0.00**
	Post-JSON	0.149	0.045	0.007	235	**18**	0.00
	Post-Rel.	0.172	0.083	0.024	214	28	2.98
Q2	MongoDB	0.431	0.079	**0.00**	151	**10**	**0.00**
	Post-JSON	0.189	0.048	**0.00**	144	**10**	3.91
	Post-Rel.	**0.097**	**0.030**	**0.00**	**220**	13	**0.00**
Q3	MongoDB	0.133	0.038	**0.00**	**558**	34	**0.00**
	Post-JSON	**0.037**	**0.022**	**0.00**	427	**22**	**0.00**
	Post-Rel.	**0.036**	**0.022**	**0.00**	453	25	**0.00**

queries, achieving 31 and 35% higher throughput than MongoDB and PostgreSQL JSON, respectively. The corresponding jitter is 0.03 s for the response time and 10–12 QPS in the throughput.

Scenario (Q3)—IIoT Correlation functions: The relational PostgreSQL performed equally to PostgreSQL JSON in terms of response time in the correlation function queries, while both of them outperformed MongoDB by almost 3.6 times. The response time jitter is of the order of 0.022 s. On the contrary, in the correlation function queries, MongoDB achieved 19 and 24% higher throughput than relational PostgreSQL and PostgreSQL JSON, respectively, with a jitter around 34 QPS.

Further Discussion: Through this experiment, some key issues of PostgreSQL were observed. In the Q1 test case scenario, a plethora of records failed to be inserted into the database, especially on the low QPS case. We suspect it occurs due to the PostgreSQL sequence number generator's inability to process simultaneous requests. A solution to this issue would be a more complex client-side fix, with the ability to resend any failed insertion operations, or more refinement on the locking system of the database, with the added side effect of slower performance. Another minor issue was observed. Performing throughput analysis on PostgreSQL utilizing the pg-stat-statements module [12] produced inconsistencies due to oversampling on its namesake database. This issue did not affect this experiment's data analysis precision although in an industrial, ultra-large-scale system has the potential to grow exponentially in severity (Table 1).

5 Conclusions

This paper presents the IIoT architecture and big data requirements for next-generation industries, such as Industry 4.0. Collecting and analyzing enormous

amounts of data require a DBMS system to be constructed, and as a consequence to be evaluated. To this direction, three different use case scenarios are presented in order to evaluate the performance of three open-source data storages, such as MongoDB, PostgreSQL with JSON, and relational Postgres. The proposed stress tests show that MongoDB outperformed Postgres with a 19–30% response time for the insert queries, achieving 51–55% higher throughput. Furthermore, relational Postgres is four times faster than MongoDB and two times faster than Postgres JSON for the selection queries, achieving 31–35% higher throughput. Moreover, both versions of Postgres performed equally in terms of response time for the correlation function queries, while both of them have outperformed MongoDB by 3.6 times. On the contrary, MongoDB achieved 19–24% higher throughput than relational Postgres and Postgres JSON, respectively, in the Q3 scenario. Finally, the authors set as future work the addition of more client nodes to the network in order to increase the maximum QPS, secondly the enlargement of the database or increase the tables' size per collection, and thirdly evaluate the performance of using BSON files.

Acknowledgements This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH—CREATE—INNOVATE (project code: T2EDK-00708).

References

1. Lasi H, Fettke P, Kemper H-G, Feld T, Hoffmann M (2014) Industry 4.0. BISE 6:239–242. https://doi.org/10.1007/s12599-014-0334-4
2. Padovano A, Longo F, Nicoletti L, Mirabelli G (2018) A digital twin based service oriented application for a 4.0 knowledge navigation in the smart factory. IFAC-PapersOnLine, 51(11):631–636. https://doi.org/10.1016/j.ifacol.2018.08.389
3. Chen B, Wan J, Shu L, Li P, Mukherjee M, Yin B (2018) Smart factory of industry 4.0: key technologies, application case, and challenges. IEEE Acc 6:6505–6519
4. Hu P (2015) A system architecture for software-defined industrial internet of things. IEEE ICUWB. https://doi.org/10.1109/icuwb.2015.7324414
5. Yue X, Cai H, Yan H, Zou C, Zhou K (2015) Cloud-assisted industrial cyber-physical systems: an insight, microprocessors and microsystems, 39. https://doi.org/10.1016/j.micpro.2015.08.013
6. Lee J, Bagheri B, Kao HA (2015) A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf Lett 3:18–23. https://doi.org/10.1016/j.mfglet.2014.12.001
7. Makris A, Tserpes K, Spiliopoulos G, Anagnostopoulos D (2019) Performance evaluation of MongoDB and PostgreSQL for spatio-temporal data. In: Workshop proceedings of the EDBT/ICDT (2019)
8. Mongodb. http://www.mongodb.com/. Accessed on 21 Nov 2021
9. Binary json. http://bsonspec.org/. Accessed on 21 Nov 2021
10. Performance Benchmark POSTGRESQL / MONGODB. https://info.enterprisedb.com/rs/069-ALB-339/images/PostgreSQL_MongoDB_Benchmark-WhitepaperFinal.pdf. Accessed on 22 Nov 2021
11. Cstore. https://citusdata.github.io/cstore_fdw/. Accessed on 26 Nov 2021
12. Postgres, pg-stat-statement. https://www.postgresql.org/docs/9.4/pgstatstatements.html. Accessed on 26 Nov 2021

13. Plugge E, Hawkins T, Membrey P (2010) The definitive guide to MongoDB: the NoSQL database for cloud and desktop computing, A Press. ISBN-10:1430230517
14. Matthew N, Stones R (2005) Beginning databases with postgresql: from novice to professional, Apress USA, ISBN: 1590594789
15. Rossman G (2021) New benchmarks show postgres dominating MongoDB in varied workloads. https://www.enterprisedb.com/news/new-benchmarks-show-postgres-dominating-mongodb-varied-workloads. Accessed on 11 Nov 2021
16. OnGres. https://ongres.com/. Accessed on 26 Nov 2021
17. Sysbench. https://github.com/akopytov/sysbench. Accessed on 26 Nov 2021
18. PostGIS. https://postgis.net/. Accessed on 26 Nov 2021
19. Klein J, Gorton I, Ernst N, Donohoe P, Pham K, Matser C (2015) Performance evaluation of NoSQL databases: a case study. Association for Computing Machinery, New York, NY, USA. 9781450333382. https://doi.org/10.1145/2694730.2694731
20. Asiminidis C, Kokkonis G, Kontogiannis S (2018) Database systems performance evaluation for IoT applications. IJDMS 10:1–14. https://doi.org/10.5121/ijdms.2018.10601

Fault-Tolerant Distributed Mutual Exclusion over Elastic Logical Ring Topology

Milen Loukantchevsky

Abstract Under *mutual exclusion* is understood preventing of any opportunity more than one active object (process, thread, task) to access a shared resource at a time. The distributed *ring-based* (aka *token-ring*) mutual exclusion algorithm is executed over logical circular topology. Due the chosen topology the ring-based algorithm is the simplest of this kind. Its pros are simplicity and minimalistic preliminary information required to be known a priori from each system process. The main drawback of this attractive algorithm in its basic definition is the strong presumption of absolute system reliability which makes it impractical. After all, the failure model of distributed systems itself assumes that failures should not be treated as exceptions but as a norm. Three working recovery schemes are considered through the overall project "Class of Fault-Tolerant Distributed Algorithms for Mutual Exclusion over Elastic Logical Ring Topology". They evolve consistently and complement each other: Failure recovery without reconfiguration (*Scheme 1*); Failure recovery with one-way reconfiguration/resiliency (*Scheme 2*); Failure recovery with two-way reconfiguration/resiliency (*Scheme 3*). The first two recovery schemes are described in previous works. Here is presented the third one. It supposes both exclusion of the faulty processes and inclusion (injection) of faultless spare processes as replacement of faulty ones. Combining both *Scheme 2* and *Scheme 3* the communication ring acquires the property of elasticity or two-way resiliency. So, we eliminate the impractical assumption of "absolute" reliability. The full code of the algorithm test bed is placed in the *GitHub* under *MIT* license.

Keywords Distributed systems · Fault-tolerance · Failure recovery · Mutual exclusion · Resiliency · Token-ring

M. Loukantchevsky (✉)
University of Ruse, 8 Studentska Street, POB 7017, Ruse, Bulgaria
e-mail: mil@ieee.org

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_26

1 Introduction

Firstly, we have the task to prevent any opportunity more than one active object (process, task, thread, etc.) to access a shared resource at a time, so called *mutual exclusion* (*ME* shortly).

Secondly, our execution environment is distributed, i.e., there are not those handy global system variables suitable for coordinating the access to the shared resources. As consequence we have only one means of inter-process communication: the message passing over physical channels. That leads to very high level of *nondeterminism* and specific *failure model* of the distributed systems: failures are not to be treated as exceptions but as a norm.

Thirdly, in the most cases we prefer as simple as possible solutions (at least at the very foundation) adhering to the well-known *KISS* principle.

The distributed ring-based (aka token-ring) mutual exclusion algorithm executed over logical circular topology is an example of such a solution which combines the above. With one, but very unpleasant exception: the requirement of absolute system reliability of its basic variant.

That is the reason to formulate our *objective*: through suitable modification to overcome the main shortcoming of the basic algorithm—the inadmissibility of process failures. The project "Class of Fault-Tolerant Distributed Algorithms for Mutual Exclusion over Elastic Logical Ring Topology" is created to achieve the objective set. Three working recovery schemes are considered through the overall project that evolve consistently and complement each other:

- Scheme 1 (Failure recovery without reconfiguration)
- Scheme 2 (Failure recovery with one-way reconfiguration/resiliency)
- Scheme 3 (Failure recovery with two-way reconfiguration/resiliency)

Under *Scheme 1* is assumed that no changes in the system configuration are made. In case of a process failure the hole system stops (at least in respect to *ME* algorithm). Only after the faulty process is restored other processes resume normal operation from the state they were in at the time of the failure, and thereafter the *ME* token is restored. In doing so, strict compliance with the *ME1* and *ME2* requirements is ensured. This modified distributed ring-based mutual exclusion algorithm designated as *Mx1ME* allows distributed system recovery from multiple faults [1].

With *Scheme 2* the system recovery begins as soon as a process failure is detected. This is at the expense of excluding of the faulty process from the system configuration. That leads to ring *shrinking* (one-way resiliency) and represents a kind of *graceful degradation* [2]. Compared to the first scheme, the second one is more operational as eliminates the time to repair of the faulty process. The whole system resumes its normal operation with latency determined by the next times: the time to detect the process failure, the time to isolate of the faulty process, the time to recover of the *ME* token.

Scheme 3 supposes both exclusion of the faulty processes and inclusion (injection) of faultless spare processes (despite if this is a recovered or a "*fresh*" one process)

as replacement of the faulty ones. Apparently, it is a kind of ring *stretching*. So, combining both *Scheme 2* and *Scheme 3* the communication ring acquires the property of *elasticity* or two-way resiliency.

2 *Scheme 3* for Two-Way Resiliency

2.1 Chain of Algorithms for Two-Way Resiliency

The modified distributed ring-based mutual exclusion algorithm, namely, *Mx1ME*, is last in the chain of helper distributed algorithms [2]. If, for simplicity here, we do not distinguish *CEH* from *CEH.GD* and, respectively, *RUP* from *RUP.GD* we got generalized form of that chain (Fig. 1). What is new in *Scheme 3* is the addition of the helper distributed algorithm *INJ* for process injection into the communication ring.

Under *injection*, implemented by the *INJ* algorithm, we understand such a change of the system configuration, suggesting addition of new process in it. It could be already recovered process, switched some time ago off from the configuration because of its failure, or a "*fresh*" one process which has not yet been part of the system.

At Table 1 we could follow the normal (without any faults) execution startup sequence within the system consisted of *3* processes. (The time format used: *ss.ms*).

Each individual startup will of course be characterized by different instants of time and will depend on number of factors, some of which under certain control (by the appropriate system constants) and other out of such control. But the general sequence of characteristic times will be as shown at Table 1, respectively, at Fig. 2.

Each helper algorithm has two characteristic times: the instant of its initiation and, respectively, the instant of its finishing. So, T_{START} is the instant of the process creation and the beginning of *CEH* algorithm, while $T_{OUT::CNC}$ is the instant of connect by

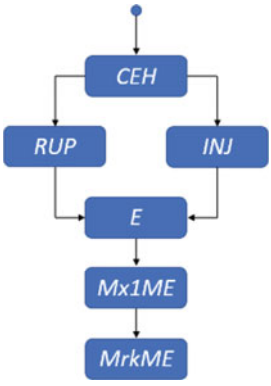


Fig. 1 Generalized chain of algorithms for two-way resiliency

Table 1 Fixing of characteristic times ($n = 3$, normal execution sample)

#	Time	P_1	P_2	P_3
0	T_{START}	00.000	01.337	02.649
1	$T_{OUT::CNC}$	03.027	04.349	05.673
2	$T_{RUP::START}$	03.029	04.355	05.677
3	$T_{RUP:UP}$	06.075	07.518	05.684
4	$T_{E::START}$	06.077	07.520	05.686
5	$T_{E::END}$	06.336	07.771	07.832
6	$T_{MrkME::START}$	07.893	07.894	07.953
7	$T_{MrkME::END}$	07.961	08.016	08.017

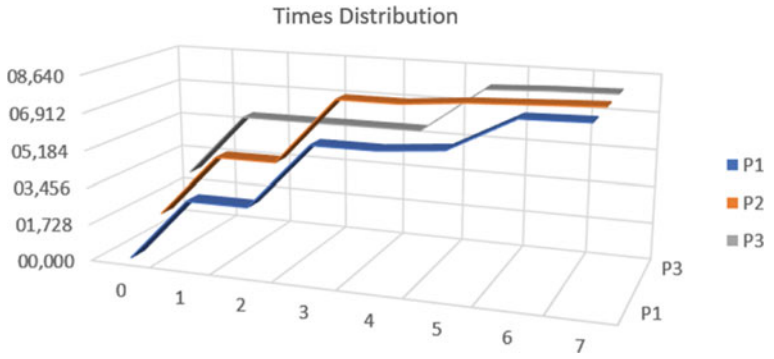


Fig. 2 Graphical representation of Table 1

output channel to the immediate neighbor; $T_{RUP::START}$ corresponds to the beginning of *RUP* algorithm, while $T_{RUP:UP}$ is the instant when the *RUP* algorithm finish with rising the event *OnRingUp*; and so on.

The research is performed through the test bed system developed and made publicly available at *GitHub* [3].

2.2 *Injection General Formalized Specification*

As already said in 2.1, under injection, implemented by the injection algorithm *INJ*, we understand such a change of the system configuration, suggesting addition of new process in it. The process injected might be one of the spare processes or a repaired process removed off the ring some time before because of its failure.

Let look at a segment of the system ring (Fig. 3). The segment contains three processes: P_{i-1}, $P_{i,}$ and P_{i+1}. (Should be noted that the ordering of the processes in

Fig. 3 Injection (opening)

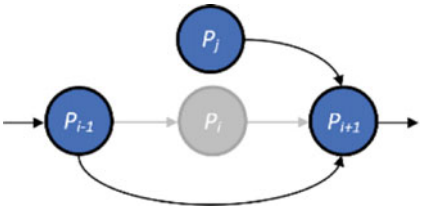
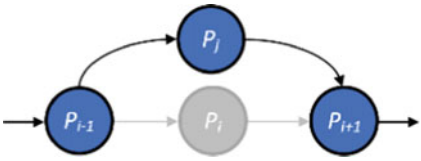


Fig. 4 Injection (closing)



the ring is irrelevant and they are arranged consecutively at the figure for convenience only.) The process P_i is currently excluded from the ring because of its failure.

The *INJ* algorithm consists of three *phases*: opening, forwarding, and closing. At the *opening* phase (Fig. 3) the injecting process P_j initiates the operation sending to P_{i+1} the message <*mrk_inj, j, k*> . The first item *mrk_inj* is the type of the message used to distinguish it from all other messages exchanged over the ring. The second item *j* is the identifier of the injected process, and the last one—*k* is the identifier of the process immediately in front of which P_j is to be placed. In our case *k* equals to $i + 1$.

At the *forwarding* phase the injection message <*mrk_inj, j, k*> is forwarded by each process until it reaches the process P_j. Furthermore, the identifier *j* is inserted into the list of system process identifiers *CEH::ListPIds* at the appropriate place next before the identifier *k*. In this phase the process P_{i-1} disconnects by output from P_{i+1} and connects to P_j.

At *closing* phase P_j receives its own injection message <*mrk_inj, j, k*> from P_j (Fig. 4). As a side effect this event corresponds to the *RUP::OnRingUp* event of the *RUP* algorithm. Hence there is not needed the newly injected process P_j to start the ring checkup algorithm. This explains why at Fig. 1 the *RUP* and the *INJ* algorithms are shown conditionally side by side.

The *INJ* algorithm is valid under next assumptions:

- The distributed system is synchronous.
- The type of process failures is *fail-stop.*
- The injected process should not duplicate a process already part of the ring.
- Failures of other processes are acceptable during the operation.
- Multiple simultaneous injections of different processes at different points are possible.

The generalized formal specification of the *INJ* algorithm consists of a *declarative* part and of definitions of *event handlers.* The declarative part includes specification of system constants, messages, states, and internal state space:

```
{SYSTEM CONSTANTS}
Int MAX_INJ_PERIOD   // period between injection attempts
PId i                // process Pi identifier

{MESSAGES}
<mrk_inj, j, k>

{SET OF STATES}
<State> := {OFF, ON}

{INTERNAL STATE SPACE}
State state         // current process INJ state
RUP::state          // current process RUP state
Timer TimerINJ      // INJ timer
CEH::ListPIds       // list of process identifiers
CEH::PIdNext        // current neighbor process identifier
```

The injection timer *TimerINJ* reinitiates the operation if it does not finish within the *MAX_INJ_PERIOD*. Thus, we ensure the fourth assumption of admissibility of other process' failures during injection. Some of the internal state space is shared with the *CEH* and the *RUP* algorithms.

The definitive part of the *INJ* of course should be consistent with the reactive character of the distributed systems—the system is in a stable state until some predefined event brought it to another stable state. That is why this part consists of the handlers of all system events caught.

```
OnInit:
   state := ON
   RUP::state := DOWN
   TimerINJ.Interval := MAX_INJ_PERIOD
   CEH::ListPIds.Set()
   CEH::PIdNext := CEH::ListPIds.Front()
   ChannelOut.Open()

OnConnect:
   {Injection First Attempt}
   OnStart()

OnStart:
   {Opening Phase}
   Send <mrk_inj, i, pidNext>
   TimerINJ.Start()

OnReceiptOf <mrk_inj, j, k>:
   If i = j
      {Closing Phase}
      TimerINJ.Stop()
      state := OFF
      RUP::state := UP
      {Distributed Election Entry Point}
      E::OnStartElection()
   Else
      {Forwarding Phase}
      If j ∉ CEH::ListPIds
         CEH::ListPIds.Insert(j, k)
      EndIf
      If CEH::pidNext == k
         {Output Channel Reconnect}
         ChannelOut.Close()
         CEH::PIdNext := CEH::ListPIds.Front()
         ChannelOut.Open()
      EndIf
      Send <mrk_inj, j, k>
      {Start Ring CheckUp Algorithm}
      RUP::OnStart()
   EndIf

OnTimer:
   {Injection Next Attempt}
   OnStart()
```

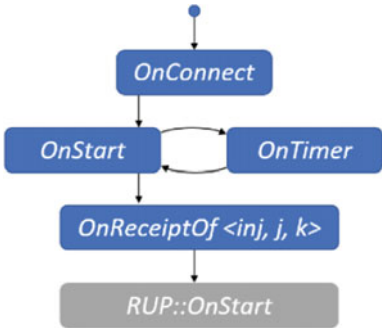


Fig. 5 INJ generalized events diagram

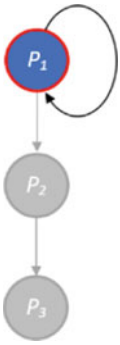


Fig. 6 INJ sample (phase 1)

Might be said that the key event is *OnReceiptOf <mrk_inj, j, k>* . Firstly, is checked if *i* equals *j*. If so, we go into the closing phase of the algorithm. Otherwise, we go into the forwarding phase.

The generalized *event diagram* (Fig. 5) represents the sequence of occurrence of system events caught. It is useful to reconstruct the inter-connection of those events.

The last event *RUP::OnStart* is distinguished from others because it does not occur in the injected process P_j but in others during the forwarding phase.

2.3 Ring Stretching Sample

Let us focus our attention on a possible scenario of execution. Suppose we have given by condition a system from three processes: P_1, P_2, and P_3. (As noted before the actual ordering of the processes in the ring is irrelevant and they are arranged consecutively for convenience only.)

Moreover, although the full system configuration includes those three processes, let assume that initially only P_1 is available, but other two are not (*Phase 1*). Initially $ListPIds_1 = \{P_2, P_3, P_1\}$. After the execution of the whole chain of algorithms the

system will reduced to its smallest possible dimension (Fig. 5). The processes P_2 and P_3 will be excluded as faulty and $ListPIds_1 = \{P_1\}$. Well, the ring consists of only one process, but this loopback ring is fully functional! In this case the election algorithm will select P_1 as coordinator—there are simply not others (The coordinator is marked in red circle).

At next phase (*Phase 2*) let try to inject the process P_3 by sending to P_1 the injection message <*mrk_inj, 3, 1*> . When it is over, we will have a ring from two processes (Fig. 7), and, respectively, $ListPIds_1 = \{P_3, P_1\}$, $ListPIds_3 = \{P_1, P_2, P_3\}$. Note that at this case the injected process P_3 does not know that P_2 is faulty, but this is not required, as every process in the ring is only responsible for its output connection. The election algorithm will select P_3 as coordinator.

Next, at *Phase 3* we try to inject the process P_2 by sending to P_3 the injection message <mrk_inj, 2, 3> . When it is over, we will have a ring from three processes (Fig. 8), and respectively, $ListPIds_1 = \{P_2, P_3, P_1\}$, $ListPIds_2 = \{P_3, P_1, P_2\}$, $ListPIds_3 = \{P_1, P_2, P_3\}$. The coordinator set by the distributed election algorithm will remain P_3.

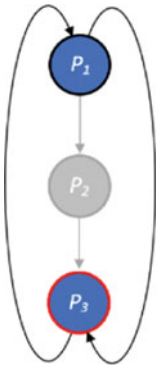


Fig. 7 INJ sample (phase 2)

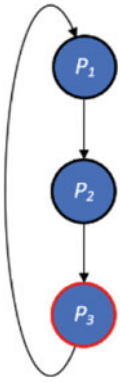


Fig. 8 INJ sample (phase 3)

Table 2 Fixing of characteristic times ($n = 3$, *INJ* sample, *Phase 3*)

#	Time	P_1	P_2	P_3
0	T_{START}	00:00.000	01:56.394	01:36.830
1	$T_{INJ::2::START}$	02:02.550	01:59.415	01:59.532
2	$T_{RUP::START}$	02:02.552	01:59.415	01:59.537
3	$T_{RUP:UP/INJ::2::END}$	02:02.570	02:02.552	02:05.695
4	$T_{E::START}$	02:02.572	02:02.553	02:05.697
5	$T_{E::END}$	02:02.578	02:05.880	02:05.944
6	$T_{MrkME::START}$	02:06.008	02:06.011	02:06.004
7	$T_{MrkME::END}$	02:06.087	02:06.143	02:06.145

For research and practical purposes is developed algorithm *test bed*. It is implemented with the *Embarcadero C++ Builder*® development environment and its *Clang*-enhanced *C++* compiler [3]. Due to the features of the socket mechanism of communications and the setup configuration provided there are two possible modes of test bed execution: Local Distributed Mode (LDM) or Real Distributed Mode (RDM). The *LDM* is used during development, while *RDM*—after deployment. Mapping from *LDM* to *RDM* is isomorphic [1].

At Table 2 we could follow the execution trace of a run of the above presented scenario's *Phase 3*. (The time format used: *mm:ss.ms*).

Here, as with Table 1, each individual run will be characterized by different instants. But the general sequence of characteristic times will be as shown at Table 2.

For the processes P_2 the characteristic time $T_{INJ::2::START}$ corresponds to the beginning of *INJ*. While, for the processes P_1 and P_3 that time corresponds to the instant of receiving of the appropriate injection message.

The characteristic time noted as $T_{RUP:UP/INJ::2::END}$ for the process P_2 corresponds to the end of injection with side effect equivalent to the time of *RUP::OnRingUp* event. While for P_1 and P_3 it corresponds to the time of *RUP::OnRingUp* event.

3 Conclusion

The subject of this article is the third of the three recovery schemes developed under the project "Class of Fault-Tolerant Distributed Algorithms for Mutual Exclusion over Elastic Logical Ring Topology". The *Scheme 3* (Failure recovery with two-way resiliency/reconfiguration) evolves consistently and complements the previous two: the *Scheme 1* (Failure recovery without reconfiguration) and the *Scheme 2* (Failure recovery with one-way resiliency/reconfiguration).

Here is shown the modification of the chain of algorithms for resiliency and the generalized formal specification of the injection.

As result we get *elastic* circular (aka ring-based) logical communication topology. It *shrinks* at process failures and *stretches* at process recovery by its injection to the current ring configuration.

The algorithm *test bed* is implemented with the *Embarcadero C++ Builder*® development environment and its *Clang*-enhanced *C++* compiler. The evolution of the test bed can be traced on the author's *GitHub* repository [3].

The implementation is focused on the exploration of the algorithm as well as on the promotion of its practical usage. All working recovery schemes proposed are incorporated and could be explored separately or jointly.

The main references for this third stage of the project are [1–3]. However, it is my duty to mention again the sources that greatly influenced the overall work:

- In the area of distributed systems as a whole [4–9].
- In the area of distributed mutual exclusion [6, 8, 10–13].
- In the area of fault-tolerance [10, 14, 15].

The author will be grateful for any recommendations that can improve the work presented. And especially to those who are interested in further development and practical use of the proposed failure recovery techniques.

References

1. Loukantchevsky M (2020) Distributed ring-based mutual exclusion with failure recovery. In: Proceedings of 21-th international conference on computer systems and technologies. ACM, New York, NY, USA, pp 111–115, ISBN 978-1-4503-7768-3. https://doi.org/10.1145/3407982.3408014
2. Loukantchevsky M (2021).Distributed ring-based mutual exclusion with graceful degradation. In: Proceedings of University of Ruse—2021, vol 60. book 3.2, pp 20–27, ISSN 2603–4123
3. Loukantchevsky M (2021) XME ring github repository. https://github.com/milphaser/XME.Ring. Retrieved 30 Nov 2021
4. Coulouris G et al (2011) Distributed systems: concepts and design, 5th edn. Addison-Wesley, Boston, p 1008
5. Kleinrock L (1985) Distributed systems. Commun ACM 28(11):1200–1213
6. Kshemkalyani AD, Singhal M (2008) Distributed computing: principles, algorithms, and systems. Cambridge University Press, Cambridge, p 736
7. Loukantchevsky M (2014) Distributed systems: theory and practice. Ruse University Press, Ruse, p 212, ISBN 978-619-7071-35-1
8. Malkhi D (2019) Concurrency: the works of Leslie Lamport. ACM, New York, NY, USA, 2020, p 366, ISBN 978-1-4503-7270-1. https://doi.org/10.1145/3335772
9. Ralston A, Reilly E, Hemmendinger D (2003) Encyclopedia of computer science. 4th edn. Wiley, p 2080, ISBN 978-0-470-86412-8
10. Banerjee S, Chrysanthis P (1996) A new token passing distributed mutual exclusion algorithm. In: Proceedings of the international conference on distributed computing systems (ICDCS). https://www.researchgate.net/publication/2666403_A_New_Token_Passing_Distributed_Mutual_Exclusion_Algorithm. Retrieved 30 Nov 2021
11. Parihar AS, Chakraborty SK (2021) Token-based approach in distributed mutual exclusion algorithms: a review and direction to future research. J Supercomput, pp 1–51. https://doi.org/10.1007/s11227-021-03802-8

12. Raynal M (1991) A simple taxonomy for distributed mutual exclusion algorithms. ACM SIGOPS Operating Syst Rev 25(2):47–50
13. Saxena PC, Rai J (2003) A survey of permission-based distributed mutual exclusion algorithms. Comput Stand Interfaces 25(2):159–181, ISSN 0920-5489. https://doi.org/10.1016/S0920-5489(02)00105-8
14. Christian C, Guerraoui R, Rodrigues, L (2011) Introduction to reliable and secure distributed programming. 2nd edn. Springer-Verlag, Berlin, p 386, ISBN 978-642152597
15. Mohammed A, Kavuri R, Upadhyaya N (2012) Fault tolerance: case study. In: Proceedings of the second international conference on computational science, engineering and information technology, Oct 2012, pp 138–144. https://doi.org/10.1145/2393216.2393240

Performance Analysis of TDM-PON Protection Schemes by Means of the PON Network Availability Evaluator

Rastislav Róka

Abstract This paper is involved in the performance analysis of protection mechanisms utilized in common time division multiplexing-based passive optical networks. First, possible protection schemes for this kind of passive optical networks involved into the performance analysis are presented. This analysis is based on corresponding reliability diagrams and parameters used for the evaluation of the total network availability for each considered protection scheme. For evaluating possible migration scenarios related to protection schemes and comparing of protection possibilities of various passive optical networks, an appropriate simulation program must be prepared. Subsequently, the PON network availability evaluator is realized for obtaining relevant results appropriate for the performance analysis. The simulation program utilizes specific parameters for particular optical components utilized in a selected protection scheme and presents its resultant network availability. Values of considered parameters are changing according to the well-known data resources or specific network operator's data. Finally, an evaluation of the total network availability for considered traffic protection schemes utilized in TDM-PON networks is included.

Keywords TDM-PON networks · Traffic protection schemes · PON network availability evaluator · Performance analysis

1 Introduction

The evolution of passive optical networks is oriented on larger geographical areas, higher numbers of subscribers and higher bandwidth per subscriber. Despite clear advantages of future passive optical access networks based on the wavelength division multiplexing (WDM) principles, many network operators permanently develop time division multiplexing-based (TDM) passive optical networks, eventually with a moderate upgrade to hybrid passive optical networks. In [1], a possible migration

R. Róka (✉)
Slovak University of Technology, Bratislava, Slovakia
e-mail: rastislav.roka@stuba.sk

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_27

scenario to the next generation of hybrid passive optical access networks is presented also with a proposal for a self-restored architecture including the classical point-to-multipoint physical topology. In [2], a hybrid passive optical network (PON) mechanism is analyzed by means of power budget and transmission performances. TDM-PON topologies can be considered with different protection mechanisms compatible with hybrid topologies [3]. In [4], traffic protection schemes considered for these HPON networks are evaluated using HPON network configurator [5–7]. Based on the used network topology, results of this evaluation are presented together with a recommendation that the TDM-PON network deployment can be done with optimized protection mechanisms for significant decreasing in capital investments. It is profitable to keep at disposal an adequate number of reserve optical fibers and/or to provide protection mechanisms at the TDM-PON network deployment. In [8], an effect of only fiber duplications on a protection feasibility of passive optical networks is analyzed. In [9], a comparison of power budget analyses for single- and dual-feeder fiber architectures of passive optical networks is presented. Various WDM-PON architectures utilizing only passive components in remote nodes are analyzed from a viewpoint of with the traffic protection securing in [10]. In [11], protection mechanisms that can be applied for the whole long-reach passive optical networks are proposed. It is suggested to consider cost-efficient protection mechanisms for TDM-PON networks as early as possible [12].

A network survivability and protection schemes are really important and necessary in advanced optical networks. From a viewpoint of network operators, a network availability is one of the most important factor that can determine a reliability and survivability of selected traffic protection schemes utilized in passive optical networks. Therefore, our performance analysis is based on MTTR and MTBF parameters that are often and regularly used for evaluation of passive optical networks in operation and that can be used for proving the feasibility of selected traffic protection schemes.

In this contribution, possible migration scenarios for deployment the TDM-PON protection schemes are introduced. Also, the aim of analyzing these protection schemes is mentioned. Concrete protection schemes involved in the performance analysis are presented together with corresponding reliability diagrams, and parameters for the evaluation are presented in Sect. 2. For analyzing various possible traffic protection schemes, the PON network availability evaluator is prepared and created (Sect. 3). In Sect. 4, an evaluation of the total network availability for considered protection schemes utilized in TDM-PON networks is included.

2 The Deployment of TDM-PON Protection Schemes

There are quite various types of protection schemes for passive optical networks defined in the ITU-T standard [13]. For developing common TDM-PON networks, not all protection schemes are considered for practical utilization by network operators. Therefore, our performance analysis is involving following cases:

- Unprotected network,
- Type A,
- Type B,
- Dual-parented Type B,
- Type C,
- Dual-parented Type C.

Nowadays, an expansion of TDM-PON protection schemes is inadequate. Their capital investments must be minimized, and simultaneously, a network resiliency must be increased sufficiently. Moreover, reserve components should transmit an extra traffic during a common operation [13]. Considering a moderate upgrade to hybrid passive optical networks, traffic protection schemes play an increasing role simultaneously in development of modern TDM-PON networks.

2.1 Reliability Diagrams for Protection Schemes

The network reliability analysis is based on reliability block diagram (RBD) representations due to their significant advantages including an accuracy, flexibility, simplicity, and visual impact [1, 8]. They allow a determination the total network availability for a given protection scheme. In the RBD scheme, each component is included as a simple block in a correlation with adjacent blocks. Connections in a given protection scheme illustrate unprotected (in series) and/or protected (parallel) components.

In Fig. 1, RBD diagrams for cases of traffic protection schemes involved into the performance analysis are presented.

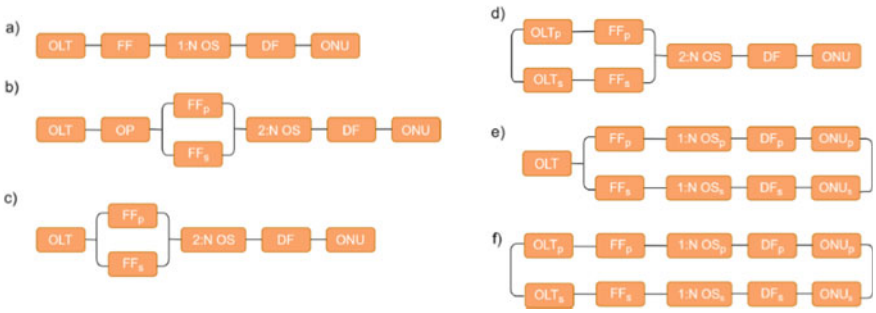


Fig. 1 RBD diagrams for **a** the unprotected network, **b** Type A, **c** Type B, **d** dual-parented Type B, **e** Type C, **f** dual-parented Type C

2.2 Parameters for the Evaluation of Protection Schemes

Each RBD block is characterized by the unavailability parameter U representing a failure probability of components [8, 13]. In a case of failure, certain network components influence a service provisioning markedly. A failure of the optical network unit (ONU) or the distribution fiber (DF) affects only one subscriber. However, a failure of the optical line terminal (OLT), the feeder fiber (FF), or the optical splitter can stop a traffic in all the passive optical network.

The unavailability of each component U_{comp} is determined mathematically as

$$U_{\text{comp}} = \text{FIT.MTTR.}10^{-9}\,[\text{h}] \tag{1}$$

where failure in time (FIT) is expressed in 10^9 h and mean time to repair (MTTR) expresses a time necessary for a failure elimination, and its value is changing for different optical components. The MTTR also depends on operational conditions. Thereby, the network availability is limited by different components in some different cases. For each component, mean time between failure (MTBF) represents a time interval among its failures and can be calculated as

$$\text{MTBF[h]} = \frac{10^{-9}}{\text{FIT}} \tag{2}$$

A relationship between the network availability A and the unavailable probability attributed to each component U_i can be expressed [13] as

$$A = 1 - \sum_{i-1}^{I} U_i \tag{3}$$

where I represents a total number of components in the passive optical network and U_i is determined as follows

$$U_i = 1 - \frac{\text{MTBF}_i}{\text{MTBF}_i + \text{MTTR}_i} \tag{4}$$

A deployment and operational situation of the real network infrastructure very differs and is depending on a specific network operator and/or area.

For evaluation of the total network availability for each considered TDM-PON protection scheme, RBD block diagrams due to their simplicity and understanding given protection schemes are used as presented in Fig. 1.

2.3 The Power Budget and a Number of Subscribers

The aim of creating the power budget (PB) in passive optical networks is ensuring a sufficient power level over the network lifetime from the OLT equipment up to the ONU equipment. The available optical budget PB [dB] can be mathematically expressed [2, 8] as follows

$$\mathrm{PB} = P_{T\min} - R_{S\min}\,[\mathrm{dB}] \tag{5}$$

where $P_{T\min}$ is the minimum transmitted power generated by the OLT laser source launched onto the optical fiber in [dBm] and $R_{S\min}$ presents a minimum average power requested by the ONU receiver in [dBm].

The signal attenuation is given as a sum of losses from all components L_{comp} as

$$\sum L_{\mathrm{comp}} = \alpha.\mathrm{FF} + \alpha.\mathrm{DF} + K.L_{\mathrm{con}} + L_{\mathrm{POS}} + \mathrm{SM}\,[\mathrm{dB}] \tag{6}$$

where FF presents a length of the feeder fiber in [km], DF presents a length of the distribution fiber in [km], α is the specific attenuation coefficient of used optical fibers in [dB/km], L_{con} is the loss due to connectors and splices in [dB], K is a number of connectors used. The loss due to a power optical splitter is represented by L_{POS} in [dB], and the system margin (SM) allocates a power penalty in the 3–6 dB range, and it is allocated during a process of network designs.

Subsequently, the calculated optical power budget $\mathrm{PB}_{\mathrm{calc}}$ in passive optical networks is determined [2] as

$$\mathrm{PB}_{\mathrm{calc}} = \mathrm{PB} - \sum L_{\mathrm{comp}}\,[\mathrm{dB}] \tag{7}$$

The power optical splitter introduces significant losses into a power budget and significantly affects a network capacity because larger bandwidth requires more reliable power transmission. Therefore, these functions must be considered at the network availability evaluation to accommodate a maximum number of subscribers for given capacity and costs in a designed passive optical network [8].

For evaluation of the total network availability for each considered TDM-PON protection scheme, a power budget and a number of subscribers due to their impacts on functionality and operation given protection schemes must be necessarily included.

3 The PON Network Availability Evaluator

The PON network availability evaluator compares possibilities for various TDM-PON traffic protection schemes. The realized simulation program evaluates real scenarios of the TDM-PON protection schemes based on various specific technical

specifications—a protection scheme, a restoration time, values of MTTR, and MTBF parameters for utilized optical components. The PON network availability evaluator is created in the Microsoft Excel program by means of ActiveX control elements and by utilization of the visual basic for applications (VBAs) code. The simulation program allows configuring specific parameters of traffic protection schemes considered for the TDM-PON networks.

The simulation program utilizes values of MTTR and MTBF parameters for particular optical components and directly presents a resultant value of the network availability for a selected protection scheme. Each component is changing independently and/or simultaneously with other components depending on the selected protection scheme. Therefore, values of considered parameters for the OLT terminal, ONU units, FF, and DF fibers are changing according to the [14, 15] or specific network operator's data, if available. The previous Table 1 gives a summary of values related to the evaluation of the total network availability for TDM-PON protection schemes.

In the PON network availability evaluator simulation program, ActiveX control elements are utilized for selecting various protection schemes and other options. Immediately, after selecting the protection scheme and setting options, resultant values of the network availability and unavailability are shown. In Fig. 2, a main interactive window of the PON network availability evaluator is presented for better understanding. By using the drop-down list box, the Type C protection scheme is selected. The restoration time is set up to 50 ms in this case. Values considered for evaluation of the network availability are used according to the data resource [15].

Table 1 Specific parameters' values for components in TDM-PON networks

Component	Parameter	Chen [14]	Hajduczenia [15]
OLT	MTTR [h] FIT MTBF [h]	4 2500 400,000	5 7000 142,857.1429
ONU	MTTR [h] FIT MTBF [h]	24 256 3,906,250	12 2500 400,000
Feeder fiber	MTTR [h] FIT MTBF [h] Fiber length [km]	6 18/km 3,086,419.753 18	14 200/km 277,777.7778 18
Distribution fiber	MTTR [h] FIT MTBF [h] Fiber length [km]	6 18/km 27,777,777.78 2	14 200/km 2,500,000 2
Optical switch	MTTR [h] FIT MTBF [h]	14 200 5,000,000	– – –

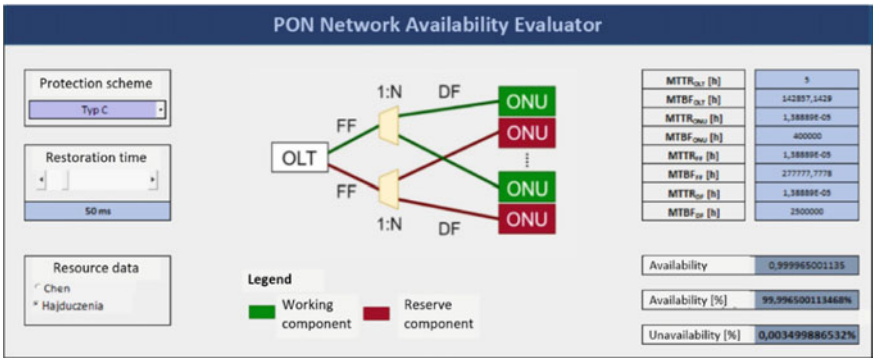


Fig. 2 PON network availability evaluator main interactive window for the Type C protection scheme

4 The Evaluation of the TDM-PON Network Availability

In Fig. 3, resultant values of the total network availability for considered TDM-PON protection schemes are presented for values from the data resource [14].

Because parameters of some components from data resources [14, 15] differ significantly, resultant graphs are also different. For comparison, a network availability of TDM-PON protection schemes at the 50 ms restoration time is presented according to data resources [14, 15] in Fig. 4. We can see that the best availability can be reached by the dual-parented Type C protection scheme where all components are redundant. From this viewpoint, it is an advantageous, however excessively cost demanding solution for designing this network protection scheme. The second-best network availability is reachable by the dual-parented Type B protection scheme. In both cases, many reserve components are utilized, above all the most important OLT terminal.

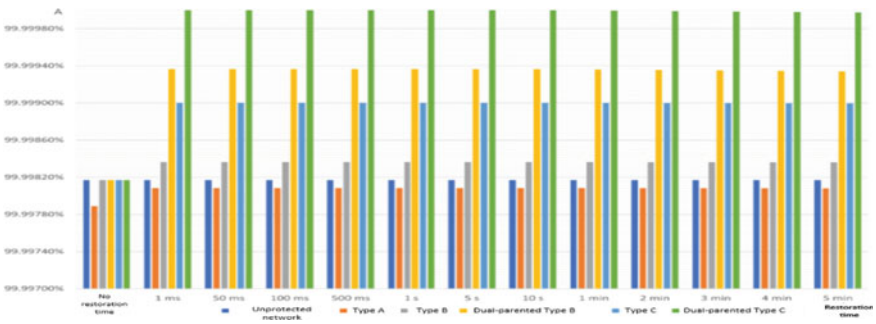


Fig. 3 Total TDM-PON network availability for selected restoration times

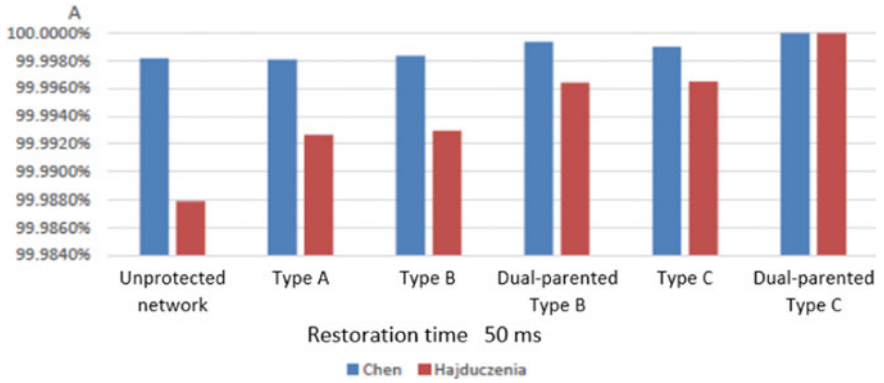


Fig. 4 Total network availability of TDM-PON protection schemes at the 50 ms restoration time for considered data resources

5 Conclusion

In this contribution, possible traffic protection schemes for common time division multiplexing-based passive optical networks are presented and analyzed. The performance analysis is based on corresponding reliability diagrams and parameters used for the evaluation of the total network availability for each considered TDM-PON protection scheme. For evaluating possible migration scenarios related to protection schemes and comparing of protection possibilities of various passive optical networks, the PON network availability evaluator is realized for obtaining relevant results appropriate for the performance analysis.

This contribution presents the first step of the research work focused on developing of future passive optical networks with new relevant technologies and related traffic protection schemes. With technological progress, new optical components and mechanisms are expecting for future passive optical networks based on dense wavelength division multiplexing. Implicitly, advanced traffic protection schemes for more reliable signal transmission and for the higher-level protection safety in next-generation passive optical networks will be considered. So, the PON network availability evaluator must be upgraded for the performance analysis of WDM-PON protection schemes in the second step of the research effort.

Acknowledgements This paper is a part of the "Utilization of Web-based Training and Learning Systems at the Development of New Educational Programs in the Area of Optical Wireless Technologies" KEGA agency project—034STU-4/2021.

References

1. Ahsan S, Lee MS, Newaz S, Asif S (2011) Migration to the next generation optical access networks using hybrid WDM/TDM-PON. J Net 6(1). https://doi.org/10.4304/jnw.6.1.18-25
2. Khan Y, Afridi MI, Khan AM, Rehman WU, Khan J (2014) Power budget analysis of colorless hybrid WDM/TDM-PON scheme using downstream DPSK and re-modulated up-stream OOK signals. J Opt Comm 35(3):231–237. https://doi.org/10.1515/joc-2013-0156
3. Mahloo M, Chen J, Wosinska L, Dixit A, Lannoo B, Colle D, Mas Machuca C (2014) Toward reliable hybrid WDM/TDM passive optical networks. IEEE Comm Mag 52(2):S14–S23. https://doi.org/10.1109/MCOM.2014.6736740
4. Róka R (2016) Optimization of traffic protection schemes for utilization in hybrid passive optical networks. Int J App Inn Eng Man 5(9):107–116. ISSN 2319-4847
5. Róka R (2014) Broadband NG-PON networks and their designing using the HPON network configurator. In: Trestian R, Muntean G-M (eds) Convergence of broadband, broadcast and cellular network technologies. IGI-Global, Hershey. https://doi.org/10.4018/978-1-4666-5978-0.ch012
6. Róka R, Čertík F (2014) Simulation tools for broadband passive optical networks. In: Pathan A, Monowar M, Khan S (eds) Simulation technologies in networking and communications: selecting the best tool for the test. CRC Press in Taylor & Francis Group, Boca Raton. https://doi.org/10.1201/b17650-16
7. Róka R (2018) Converged Fi-Wi passive optical networks and their designing using the HPON network configurator. In: Trestian R, Muntean G-M (eds) Paving the way for 5G through the convergence of wireless systems. IGI Global, Hershey. https://doi.org/10.4018/978-1-5225-7570-2.ch006
8. Imtiaz WA, Iqbal J, Qamar A, Ali H, Idrus S (2017) Impact of fiber duplication on protection architectures feasibility for passive optical networks. In: Róka R (ed) Optical fiber and wireless communications. IntechOpen, London. https://doi.org/10.5772/intechopen.68237
9. Imtiaz WA, Khan Y, Qamar A, Khan J, Khan NA (2014) Power budget analysis of dual/single feeder fiber WDMPON. Opt Lett 10(2):137–139. https://doi.org/10.1007/s11801-014-3169-9
10. Róka R (2018) Optimization of the optical power budget for various WDM-PON network designs with traffic protection securing. In: 26th International conference on software, telecommunications and computer networks. IEEE, Split, pp 76–81. https://doi.org/10.23919/SOFTCOM.2018.8555824
11. Zhang S, Ji W, Li X, Huang K, Yan Z (2016) Efficient and reliable protection mechanism in long-reach PON. J Opt Comm Net 8(1):23–32. https://doi.org/10.1364/JOCN.8.000023
12. Mas Machuca C, Chen J, Wosinska L (2012) Cost-efficient protection in TDM PONs. IEEE Comm Mag 50(8):110–117. https://doi.org/10.1109/MCOM.2012.6257535
13. ITU-T G-series Rec.—Suppl. 51 (2016) Passive optical network protection considerations
14. Chen J, Wosinska L, Mas Machuca C, Jaeger M (2010) Cost versus reliability performance study of fiber access network architectures. IEEE Comm Mag 48(2):56–65. https://doi.org/10.1109/MCOM.2010.5402664
15. Hajduczenia M, Chengbin S, Zhen Z, Elbakoury H, Kozaki S, Matsuoka M (2012) Resilience and service protection for ethernet passive optical networks in SIEPON. IEEE Comm Mag 50(9):118–126. https://doi.org/10.1109/MCOM.2012.6295721

Using Case-Based Reasoning in System Diagnostics and Maintenance

Neyko Neykov and Svetlana Stefanova

Abstract Taking optimal decisions in many cases is related to knowledge, acquired in a relevant domain. The amount of information that forms that knowledge directly affects the quality of decisions. Neither large nor too small amounts of information lead to perfect problem solutions in areas where knowledge is particularly important. The use of case-based reasoning systems can help improve performance in situations where such knowledge is scarce or non-existent at all. Our goal is to offer a model of a system that will help diagnose and solve problems met by users when they use technical products, consisting of many software and hardware components.

Keywords Knowledge management · Case-based reasoning · System troubleshooting

1 Introduction

Companies manufacturing high-tech products typically provide diagnostic and troubleshooting services for customers. It is normal for end-users to meet difficulties when using such products, and companies often build maintenance departments involved in problem's resolution. Qualification of the personnel hired in these departments affects problem's escalation to higher and more expensive levels, and this influences enterprise productivity. Today's labor market is dynamic, and skilled workers often change their positions. At the same time, the process of building competencies is slow. This impacts the profitability of companies. They have different strategies to compensate for these imbalances, and implicating systems for knowledge application can greatly increase their efficiency.

Knowledge application is one of the main processes in knowledge management. In enterprises where knowledge utilization is intensive, its dissemination and

N. Neykov (✉) · S. Stefanova
Angel Kanchev, University of Ruse, Ruse, Bulgaria
e-mail: nneykov@uni-ruse.bg

S. Stefanova
e-mail: sstefanova@ecs.uni-ruse.bg

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_28

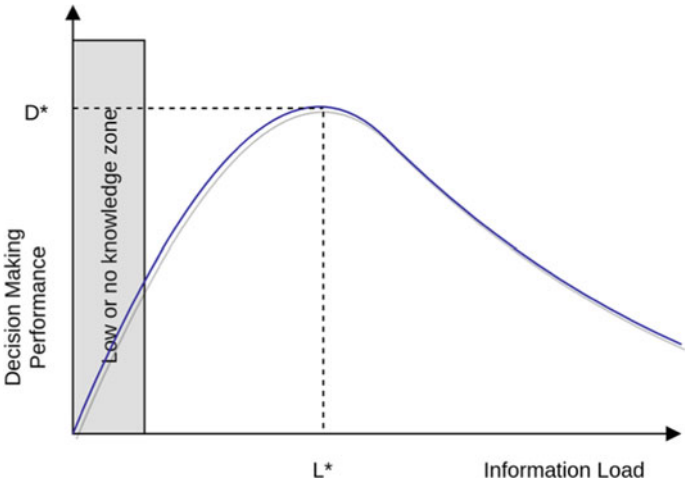


Fig. 1 Decision-making performance dependency from information load

application are of particular importance. The goal of this article is to propose a model for building a diagnosing and troubleshooting system, which helps operators support customers in using technical items. Many regular customers have difficulty using complex products and need suitable support. This could keep them as company customers in conditions of growing competition. Their expectations from the company's support department are to receive adequate advice on operating these products. Only well-trained staff can provide proper maintenance, and it is built over a long time period. However, companies do not always succeed in educating people to do this successfully. In such an environment, it is inevitable to use special artificial intelligence agents to support the application of accumulated knowledge.

Figure 1 shows the well-known relationship between the information load and decision quality. Our interest is the gray area where initial knowledge acquisition is applied, and either very little or no knowledge is available at all. Possibilities to use and apply knowledge and experience gathered by others can improve the results in such moments. Further, we examine the types of systems for collected knowledge application. We then analyze the capabilities of case-based reasoning (CBR) systems and provide a model for implementation, which includes data presentation and methods for searching, adapting, and applying solutions in such a system.

2 Knowledge Application Systems

Knowledge application processes and systems are used by individuals to apply knowledge gained by other individuals without even acquiring that knowledge [1].

Simultaneous use of knowledge management mechanisms and knowledge technologies can help build such systems. The most popular, usually in the field of artificial intelligence (AI), applicable in such intelligent agents' development are

Rule-based systems (RBS): In these systems, knowledge is represented as decision trees, and prior knowledge is required in the area to implement such a system. The set of rules and descriptions should be finite and manageable.

Case-based reasoning (CBR): processes are applicable in areas for which we have no prior or sufficient knowledge. They do not require initial full and comprehensive domain knowledge. The information base for solutions can be expanded, improved, and completed with various new cases over time. CBR expert systems can be used to train new or inexperienced professionals [2].

Constraint-based reasoning: Sometimes, there are areas where knowledge is about what cannot and should not be done. Using this approach, the problem is modeled as a hypothesis with some limitations in the conclusions, and its solution is based on satisfying these constraints.

Model-based reasoning: In the field of artificial intelligence, model-based reasoning is a method used in expert systems built on models from the physical world and descriptions of engineered system internal workings. This knowledge is typically available from design specifications, drawings, and books. Such systems can usually be used to recognize and diagnose system malfunctions [3].

Construction of expert knowledge systems involves two types of specialists—domain specialists and knowledge engineers. Cost of development is highly related to the involvement of these resources in a project and can increase if modifications need to be applied constantly to the system.

Rule-based reasoning has both advantages and disadvantages. It has widespread use in knowledge application systems. Domain knowledge specialists and knowledge engineers work together to achieve functionality in such an expert system using heuristics and rules of the type IF–THEN. This is the most popular paradigm for presenting knowledge. The 'IF' part contains conditions or prerequisites, and it serves to verify the true state of many statements. The 'THEN' part contains facts, consequences, conclusions, or actions feasible and should apply due to prerequisite certified truth [1]. Building rule bases can also be automated to some extent using machine learning techniques.

Rule-based reasoning systems also have shortcomings, some of which are critical [3]. When the amount of knowledge that must be presented is substantial, it becomes hard to maintain and validate all the rules in the base. It is difficult for rule-based expert systems to solve problems that are not explicitly covered and described. Adding new knowledge to the case base requires more effort, as domain specialists cannot make rule base extensions alone. Knowledge engineers also need to be involved in updates, which in turn makes the system more expensive, as many resources need to be invested in maintaining the system.

In its classic form, the customer support department is divided into three levels [4]. First level usually has direct contact with customers and processes complaints or answers questions. This level is not necessarily involving experts, but rather people with good communication skills. The goal here is to collect as much information as

possible about the problem, and if simple enough and within personal competencies, to handle it. Work at this level should be limited within a short period. If there is no possibility to solve the problem, it is handed over to the next level.

At the second level, domain specialists and experts are involved. They analyze problems in detail, provided the information collected at the first level. If the problem is solved, the solution is returned to the initial level and then to the client. Additional communication cycles are possible to gather more information if it is not enough. Failure to solve the problem at this level is usually due to reasons beyond control of the department or organization, and this causes problem escalation to the next level.

The tasks at third level are similar to those at second, except that experts are outside the department or company. With the assistance of these specialists, changes in software and hardware components that users operate with are possible, and solutions returned to the first level include these changed components.

In their work, different levels can actively use knowledge sharing systems, such as customer communication applications, task management applications, corporate wikis, and others.

3 CBR Model

CBR applications development requires time and specific skills, in addition to programming. There are some libraries and visual design tools developed and available to support the process. The most popular frameworks are myCBR Workbench and COLIBRI Studio, which provide user-friendly environments based on the Eclipse IDE, and can be used to design CBR systems [5]. myCBR Workbench includes tools for generating case representations directly from raw data, a graphical interface for modeling similarity metrics between specific case attributes and search quality assessment capabilities. COLIBRI is a platform for complete CBR solutions development. It provides an environment where users can share decisions and collaboratively build an expert system based on different case types, like text cases and those that use knowledge or data intensively. Case maintenance and visualization tools are also very handy.

These libraries provide opportunities to build CBR expert systems, but they are general purpose frameworks and do not provide a methodology for building a system in a specific domain. They are convenient for research and CBR practice study, as well as for small projects, but not for larger industrial applications. Our goal is to model expert systems in a specific area—the maintenance of complex products consisting of many hardware and software components, where each or all of them could give defects, or their setup and use may need additional help.

We aim to describe a model for creating an expert system based on case-based reasoning, which will help diagnose and solve problems arising while working with these products and services. In this regard, we will summarize a methodology for selecting data models, as well as work process organization, to assist the support team members in making the right decisions in their work.

For modeling case data, we choose a mixed approach and a combination of structured and unstructured elements. It is difficult to represent observed symptoms of defects and problems with structured data, and it is better to use a free descriptive form, containing text, images, references, etc. The same assumptions can be made for solution description in each case. Descriptive approach for solutions is more appropriate in this case as it gives freedom in defining a series of actions that must be taken to eliminate the defect. The role of links used in descriptions is to integrate cases with other knowledge sharing systems. New knowledge acquisition during its application can significantly increase people's qualification, opposed to a process in which only pure application takes place without awareness of the knowledge used in the case definitions.

For some data, it is possible to use a structured approach. For some parts of the observed system, like products and components they are assembled of, we may have full knowledge. This means we know this data model, so we can use structures and a more formal object schema to describe them. Using ontologies to define and store such data can be useful, because they provide opportunities to assess semantic similarity in searches.

Search for a case solution is performed by finding the nearest items in the case base stored in the system's memory. This memory is not necessarily a RAM, especially when it comes to large databases for analysis. Mixed data approach in modeling makes relational databases or other structured repository usage difficult and impractical. Therefore, NoSQL databases, unstructured repository services, or a mixture of them can facilitate the development process and improve productivity.

Several main objects can be identified when modeling the CBR expert system to support help-desk activities. These are—users, with their roles, rights, and cases, containing problem's description and its solution. Components, environment, and location, where the reported issue is observed, can be used as additional auxiliary data, giving a clearer understanding of the subject.

Maintaining information about users and their rights is important, because this way we can determine the activity of process participants and their expertise. Users can search in the system, give estimates when applying described solutions, and adapt relevant cases. On the other hand, users with more expert knowledge can modify existing cases or create completely new ones in the knowledge base, as well as review and confirm information added by other less experienced users. Performance assessment would play an important role in the knowledge base development and expert system quality.

Cases are containers for aggregating information. They define several collections—symptoms, participating components, environment, and steps to apply the solution. They are the main subject of searches in the expert system. A case is a complete problem description and provides all knowledge necessary to solve it.

Symptoms are external manifestations of the studied system, which are observed by users and participants in the troubleshooting process. Symptom explanation is difficult to present in a structured form, as it may comprise different data types like free text, images, video, etc. The freedom of expressing information in this description is important for both end-users and problem analysts.

Components that build up a complex system are perfect for structured representation. Ontologies could be a good solution for describing such a structure. Semantic networks can be used to define the interactions and connections between them. This knowledge description can help solve problems but can also be reused in other knowledge management systems.

Location and environment data are closely related to physical location, hardware, and software environment where the defect symptoms are observed, if such information is specified by the customer or support personnel. Location of the problem and the environment are not critical in solving it, but can be used to associate with similar problems elsewhere, as well as provide additional information about the environment in which system components operate (Fig. 2).

Steps for implementing the solution are defined as a series of actions that the operator must take to solve the problem. Information here is also impossible to cover with data structure, and it is preferable to use free text combined with other data formats like multimedia, etc.

In each case, links to other knowledge sharing systems can be added as an integration. With such binding, we can expect

1. Better justification of case descriptions and solutions. The staff who use the system can gain more knowledge, not just how to apply the described case solution steps without going into details. This understanding of details is not necessary to solve the case but provides opportunities for training and improving competencies.
2. Increasing the opportunities for case adaptation.
3. Improving the assessment of case quality. When the description and solution of the case contain references to shared knowledge, solution effectiveness can be estimated with some approximation before its application.

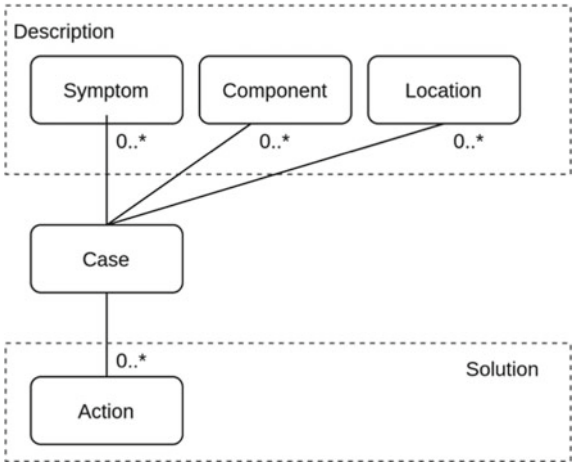


Fig. 2 Case data model

Integration with other knowledge management systems increases adaptability for new cases. Available shared knowledge can be successfully applied to adapt them, as well as to validate existing ones.

A search for cases is performed by submitting a query request from the user. The query must have a structure similar to the case description, and only a part or logical combinations of elements may be used. The calculated proximity of the data defined in query and stored cases is defined as a combination of individual element similarities participating with a certain weight in overall assessment. Use of predefined similarity weights for case attributes is convenient in a search, because it does not force users to choose one. For the more advanced, it is appropriate to further adjust weights when comparing attributes dynamically, to achieve additional flexibility in a search.

In most CBR applications, the K-nearest neighbor (KNN) algorithm is used for searching cases close to those specified in the query. As data used for attributes are heterogeneous, it is impossible to use one and the same metric for all. Instead, a combination of different metrics to calculate the total distance, assuming weights for attributes participating in the query, should be applied. Both semantic similarities and vector space model statistical algorithms are involved in searching unstructured texts. Structured data semantic similarity can be calculated relatively quickly and easily, but unstructured texts need to be pre-processed, which complicates analysis.

Reuse of cases as described in the database or by ignoring certain differences is a trivial practice, but for some of them differences must be taken into account, and some adaptation should be applied to match the real case specified in the query [6]. There are basically two ways to make this adaptation: (a) reuse the open solution by applying multiple operations to transform its attributes to the real conditions or (b) reuse the existing method of decision using the existing solution plan, but in a new context.

The process of adaptation is knowledge intensive. Some authors claim adaptive knowledge can be retrieved within a case base [7]. Domains which intensively use knowledge and systems that change in time need more knowledge than the one hidden in the case base. Providing integration with knowledge sharing systems can help and improve adaptation.

Each system is dynamic, evolves in time, and changes its functionalities and components. The rate of these changes may be different. This means that problem-solving cases can become either inaccurate in describing symptoms and solutions or may be obsolete at some stage in the system's lifetime. In order to maintain case base information correct and up-to-date, it is necessary to establish a continuous validation and verification cycle, based on user assessments or statistics for usability and quality of cases.

Performance measures are an important part of a system's development and evolution. CBR system evaluation can be sought in two directions—case resolution time, including the technical times for case detection and adaptation, and the probability of case resolution [8]. The likelihood of resolving a case depends on the quality, completeness, and scope of individual cases. Existing knowledge integration additionally can increase the adaptability.

4 Conclusion

The presented model for CBR expert system implementation is closely related to a specific domain—namely technical support provisioning in software and hardware products usage. It cannot (or it is difficult to) be reused in other areas without further development, unlike libraries such as jCOLIBRI and myCBR, which aim for reusability. By integrating knowledge application with knowledge sharing system, we want to provide additional case-specific information and assist staff with little or no expertise in adaptation and decision-making. Combination of knowledge and its practical implementation can also help the internalization process or conversion of explicit to tacit knowledge and boost the effect of gaining and understanding new knowledge. This knowledge is of practical importance when adapting existing cases or describing new ones.

As a further development, we can continue with this model application to support developers in cases where they have difficulties with the initial environment configuration and code base usage for the projects, they are working on. In many cases, even experienced developers sometimes find it challenging to set up a completely new and unfamiliar project.

References

1. Becerra-Fernandez I, Sabherwal R (2010) Knowledge management systems and processes. Sharpe Inc., M.E
2. Göker M, Howlett RJ, Price JE (2004) Case-based reasoning for diagnosis applications. In: Knowledge engineering review, vol 20, pp 117–125
3. Pantic M (2005) Introduction to machine learning and case-based reasoning. London
4. Roth-Berghofer TR (2012) Learning from HOMER, a case-based help desk support system. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 3096, pp 88–97
5. Roth-Berghofer T, Recio-García JA, Severing-Sauer C, Althoff KD, Diaz-Agudo B (2012) Building case-based reasoning applications with myCBR and COLIBRI studio. In: 17th UK workshop on case based reasoning, vol D, pp 71–82
6. Aamodt A, Plaza E (1994) Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun 7(1):39–59
7. Hanney K, Keane MT (1997) The adaptation knowledge bottleneck: how to ease it by learning from cases. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 1266, pp 359–370
8. Acorn TL (1992) SMART: support management automated reasoning technology for compaq customer service. In: IAAI-92 proceedings

Dr. AI: A Heterogeneous Clinical Decision Support System for Personalised Health Care

Md. Samiullah, Pankaj Chandra Kar, Md. Sahidul Islam, Md. Tanvir Alam, and Chowdhury Farhan Ahmed

Abstract Doctors and health workers are the key front-liners fighting relentlessly to save human lives from various diseases and hidden enemies, such as germs and microbes, since the dawn of civilisation. However, extreme work pressure due to colossal number of patient with finger-countable health workers and doctors makes it hard to maintain the quality of their service. Medical science is a vast, incredibly complicated, and ever-growing field that make the task of health workers harder and urge them to remain updated always. Hence, an AI assistant to suggest possible diagnosis and treatment options can ease the life of doctors to some extent and keep them updated. Existing clinical decision support systems (CDSSs) are either knowledge-based or pattern-based, and very few of them are for personalised care. In this paper, we propose and develop (a prototype version of) a heterogeneous CDSS using pattern mining, machine learning, and knowledge-based techniques. Finally, we conclude the paper with scopes for future extension of the proposed system.

Keywords EHR · CDSS · Pattern mining · Machine learning

1 Introduction

Public health is one of the main sectors that every country prioritises. There are immense scopes to deploy cutting-edge technologies in this sector. In this modern age, life becomes easy through the applications of modern technologies such as artificial intelligence (AI), machine learning (ML), and data/pattern mining (DM). From the dawn of modern era, the world has witnessed the wave of transformation

Supported by the Department of ICT, Ministry of PTIT, Bangladesh.

Md. Samiullah (✉) · P. C. Kar · Md. S. Islam · Md. T. Alam · C. F. Ahmed
University of Dhaka, Dhaka 1000, Bangladesh
e-mail: samiullah@du.ac.bd

C. F. Ahmed
e-mail: farhan@du.ac.bd

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_29

from the traditional system to the digital system in various sectors, including the health sector. This transformation makes a large amount of digital data available. To best use the available data, data-driven automation is a great option. Various data-driven intelligent systems have been proposed and developed which are being utilised in many real-life applications.

In medical science, most of the underlying factors and associated events involved with the factors of diagnosis and treatment are uncertain (i.e. not guaranteed). To deal with such uncertainty and inter-dependency of factors, we need probabilistic analysis, correlation analysis between the underlying factors and search for hidden factors.

CDSS is computerised systems developed to support clinicians in making decisions whilst diagnosing patients [4]. CDSS provides medical personnels, such as physicians, hospital staffs, patients, or relevant people with knowledge extracted intelligently to escalate health care. The knowledge can be used to enhance decision-making in clinical process including clinical guidelines, patient reports and summaries, and diagnostic support. Current CDSSs rely on knowledge-base constructed through a labour-intensive method or derived automatically using pattern mining approaches. The primary goal of modern CDSS is to assist medical practitioners in providing care to their patients by helping in analysing the complaints of the patient and diagnosing, based on patient data [2]. Unlike early CDSSs, those were conceived as decision-maker for medical practitioners, modern CDSS is used by medical practitioners as a source of recommendation/direction towards diagnosis in association with their own knowledge, experience, and expertise. Ideally, a CDSS suggests to guide the clinicians, and the clinician will pick the suggestion if they find it useful otherwise disregard that [2]. A clinician can decide on a set of potential diagnoses which ones are relevant and which ones are not. Then, if requires to narrow down the diagnosis, they may suggest further tests.

CDSS is categorised into classes based on: (i) knowledge extraction and processing method, and (ii) timing of use. The first kind of classification consists of two main types: knowledge-based and pattern-based [2]. On the other hand, based on timing of use, CDSS classes are pre-diagnosis, during diagnosis and post diagnosis.

A knowledge-based CDSS, typically, consists of three parts: (i) a knowledge base, (ii) an inference system, and (iii) an interface. The knowledge base is a collection of rules formed by the mined associations of data [2]. There might have facilities to update the knowledge base to meet the requirement of the system. The inference system helps in reasoning in the form of recommendation of potential diagnosis. It draws connections between the rules from the knowledge base and patient's health data. The interface is for communicating the outcome of the inference system and the input from the outer world.

Pattern-based CDSSs use machine learning approaches that allow the system to learn from past experiences in the form of extracted patterns from health and medical data. However, since machine learning systems follow mathematical models built upon the patterns of the previous diagnosis, the explain-ability and accountability for their conclusions are poor. Apparently, clinicians are reluctant in using or replying on them directly for diagnoses. A significant difference between the two abovemen-

tioned CDSSs is that the latter focus on a narrow list of symptoms whereas the former covers a wide range of symptoms involving the diagnosis of various diseases.

As per the other classification of CDSSs, pre-diagnosis CDSS assists doctors to prepare the diagnoses. CDSS those are used during diagnosis is capable to refine and narrow down the physician's preliminary diagnostic choices to ensure better outcome. Post-diagnosis CDSS mines data to capture the correlation between patient data and medical research to predict future events.

Electronic health record (EHR) is a representation of real-time patient health data to support high-quality care. It is believed that the incorporation of EHR and CDSS together in clinical diagnosis system has the capability to impact the way of treatment being taught and practised [9].

2 Dr. AI: The Proposed Platform

The proposed platform, 'Dr. AI', is a trained, intelligent system to assist doctors in predicting diseases and provide personalised treatment based on patients information. The system will allow the doctors suggesting different investigations and diagnose a patient with the help of the recommendations provided by the system.

2.1 Challenges

Incorporation of EHR in CDSS is at the same time essential and challenging. Being an uncharted area as well as having underlying complications of both deployment and implementation phases are two of the root of the challenges. Nevertheless, EHR implementation has received great attention. The implementation in health care sector faces the following challenges: (i) maintaining efficiency and safety during and after deployment [6], (ii) understanding the users' perspective of the EHR and CDSS is crucial to the success [8], and (iii) adoption of changes needs to be actively propagated [10].

Integrating new data into the existing CDSS is not only laborious but also a hard problem to deal with. Moreover, CDSSs face steep technical challenges because biological systems are extremely complicated with a vast range of interrelated data. For example, a computerised clinical system may potentially consider a patient's symptoms, medical history, family history, and genetics, as well as historical and geographical trends of disease occurrence, and published clinical knowledge when recommending a patient's course of treatment. Apart from numerous clinical researches and trials are being published and some of them are mutually conflicting. Incorporating such enormous and conflicting publications in the CDSS is quite a big challenge.

Gathering the clinical data, transforming them to medical knowledge and representing them in computer understandable format is yet to be matured. A significant amount of efforts needed to create feasible techniques to gather, store, and process

clinical data along with finding suitable, scalable, and flexible solutions to mine patterns and discover the inter-correlations between the data and the patterns. Although a significant number of efforts have been taken to build feasible CDSSs to support all kinds of clinical tasks, very few got success. On contrary, some CDSSs with limited scope have gained success, such as the leeds abdominal pain system, on behalf of the University of Leeds hospital, produced 91.8% correct diagnosis where clinicians' success rate was 79.6%.

Finally, whilst deploying the CDSS and in drawing attention of the doctors, gaining their trust on our system, and motivating them to engage in improving the system.

2.2 Developing the System

Our objective is to build an intelligent system that will be able to perform diagnosis of diseases given the symptoms or other factors such as the lifestyle of the patient and disease history. Once data are collected from various sources, a heterogeneous decision support algorithm will be developed. When the system is built, it can be used to perform various reasoning and decision-making to suggest personalised treatment. After the system development phase, the mechanism and workflow of the system will be consulted with and verified by expert clinicians. Finally, the decision support system will be made available through a Web application to the health staffs.

The proposed system works as the core/brain of the system and run in the back-end in a server. There is a web-based application for end users, such as doctors to use the system and generate prescriptions, diagnose a patient, and enhance the knowledge base of the system. The system is composed of three main components, namely (i) EHR system, (ii) knowledge base, and (iii) inference engine with self-learning algorithm.

An EHR is one of the essential part of the CDSS system. It consists of the user interface and the database schema. The interface is intended to collect patient information in an efficient and timely manner, and the database schema is targeted at storing the information in a well-organised manner to facilitate efficient processing in constructing the knowledge base. Detail of the two parts of EHR is beyond the scope of this paper. The knowledge base is constructed and enhanced in two ways: (i) pattern mining and (ii) knowledge from books, Websites, and experience of clinicians. A brief discussion of the knowledge-base construction is provided later. The final and the most important part of the CDSS is the inference engine. It takes into account the rules in the knowledge base and the newly added patient information. Based on the similarity between the new information and the rules; the trend of diseases, health, medication, and diagnosis; a reasoning is performed to recommend best potential diagnosis and treatment.

The major steps of our approach towards the development of the system are (i) data collection, (ii) knowledge-base construction, (iii) model building and training, and (iv) prototype implementation and deployment of the system.

Knowledge-base Construction The collected data require some special of natural language processing and text mining to convert them into rules, the prime components of the knowledge base. To cope up with continuous acquisition of data and best utilise it, the knowledge-base construction is made incremental. Whenever a new data, such as new patient information, new piece of diagnosis, and new publication with new findings are available, the knowledge base is updated accordingly.

In the core of the CDSS in constructing the knowledge base, there are

1. Pattern mining algorithms to find patterns and rules.
2. Text mining, Web crawling, and natural language processing (NLP) techniques to contexualize the contents of publications and building rules.
3. Transfer learning algorithms to adjust rules and correlation amongst the rules in order to cope up with spatio-temporal variations of diseases.
4. Machine learning algorithms to enable the system to learn and self-tune from the generated rules and the experience of previous predictions of diagnosis.

Model Building and Training The flow chart in Fig. 1 demonstrates the architecture and working model of the proposed system. It shows the overall procedure of collecting data, building knowledge base, and inference engine, self-tuning the model based on the previous recommendations/predictions, in a nut-shell.

We mine association rules from frequent and sequential patterns from the collected data. Because, frequent patterns provide the most practised and trendy diagnosis, whereas sequential patterns capture the order of incidents of the events that is must to be considered. Note, some data need adjustment to follow the spatio-temporal variations. To cope up with this, transfer learning approach can be used. Trend analysis in health sector can capture various pandemic and outbreaks that need special care. Furthermore, some trends in medication and treatments should be followed globally which can be captured by trend analysis. Here, time-series analysis can be performed to find trends. Data gathered through the integrated user interface are also transformed directly into rules using the aforementioned pattern mining algorithms.

Another source of generating rules for constructing knowledge base is experts' opinions obtained from books, publications, case studies, and opinions. Extracting rules in this method is laborious and time consuming. Hence, we devised an approximate algorithm to automate the rule extraction from texts where various text mining techniques, such as word embedding, word association, and tagging, have been used.

All the rules extracted using the abovementioned methods are then assigned weight to reflect the importance of the rules based on the source of the data from which the rules were extracted. Initially, in our proposed system, we assigned equal weight to the rules which can be changed as per requirement, e.g. experts' opinion-based rules should be given higher weight than the pattern-based rules.

The weights of the rules are considered whilst providing any clinical decision or recommendation. If a new patient or disease information is given to the system, the information is transformed into EHR format, and then similarly as case-based reasoning (CBR) [1], the most likely diagnosis, i.e. the diagnosis having maximum similarity score, with highest weight of the associated rules is suggested. In case of multiple diagnosis options, the options are sorted in decreasing order of similarity and weights.

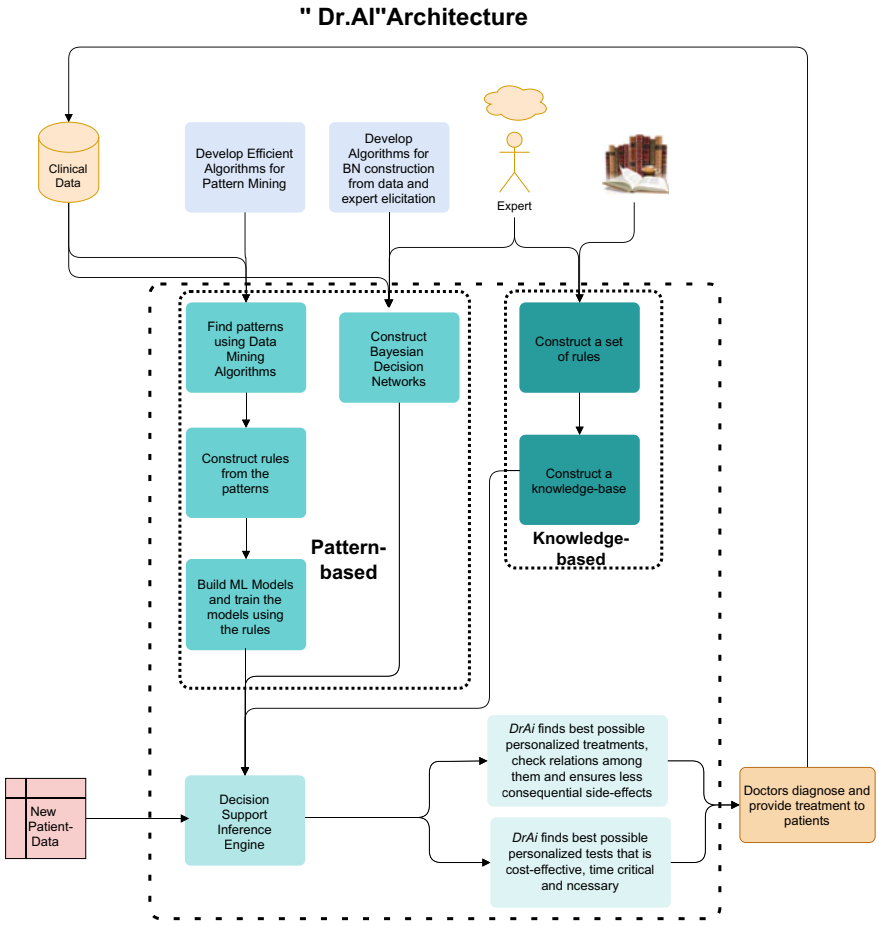


Fig. 1 Project development model for *Dr. AI*

Once a diagnosis is chosen, the complete EHR for the new patient data is added in the database from where new patterns and rules might be generated or weights of the existing rules might be adjusted. Through this process, the CDSS keeps improving itself and provides better recommendations.

The weight adjustment of the rules and weight-based ordering of possible recommendations are performed within the inference engine of the model. It is the core part of the CDSS with various decision-making and risk analysis technologies integrated in it. Initially, we are using Bayesian decision networks to choose the best diagnosis using utility and decision theory. It is quite challenging and laborious job to construct BDNs for a vast and overly complicated field, medical science. One option is to use automated construction of BDNs which is still in its infancy.

Prototype Implementation and Deployment of the System We are currently developing a prototype version of the proposed model for the CDSS and planning to deploy it as a Web-based application. A non-exhausted list of concerns in developing and deploying such a CDSS is [2]: (a) privacy and confidentiality, (b) user-friendly interface and access for all class of people, (c) accuracy and completeness of information, (d) integration of new information and requirements, (e) compatibility with the existing systems, (f) uniformity and non-conflicting knowledge, and (g) acceptance by the core-community.

In developing the prototype of the model, we considered all the aforementioned factors. We have been using Python 3.5 as the programming platform and Flask Web development framework of Python. The system contains a user-friendly interface for doctors, patients, other health workers, and relevant people with proper access-right checking facility in order to ensure privacy and confidentiality. A doctor can generate a prescription with the least effort and time effective manner. This makes the application widely acceptable in the community. We provide create-retrieve-update-delete (CRUD) facilities for information, such as prescription, test reports, and health records. Patients, now, do not need to inform the clinician about their past disease history, allergies to medications, and other health-related information as the doctors can find them from the stored EHR for the patient.

The developed prototype is designed to be extended easily. It is a plug-in-based system where new features can be easily added and any piece of code, such as a particular pattern mining algorithm, time-series approach, or machine learning technique can be easily replaced with the interface of the code segment preserved.

3 Experimental Analysis

As a proof-of-concept, we have conducted experiments using a bag of classifiers. The experiments were conducted on a set of 471 real medical prescriptions having 2006 distinct items and a medical transcriptions [3] dataset, comprised of a set of 3823 sample medical prescriptions with 11,525 distinct items.

In case of the first dataset, the classifiers that were used with their accuracy: Multi-layer perceptron classifier (neural network): 61.74%, support vector machine: 30.67%, decision tree classifier: 42.19%, logistic regression classifier: 56.79%, and an ensemble classifier of the first four classifiers with max-voting strategy: 58.02%.

In case of the second dataset, the number of distinct labels found is huge that refrains other algorithms to be executed in general purpose computers. Hence, a multi-layer perceptron classifier model is built for this dataset and the average accuracy obtained by the model is 83.57%.

4 Conclusions and Future Works

Health care is a complex, vast, and sensitive field where occurring errors are a common phenomena. Apparently, the effort to minimise them is important [5, 11]. CDSS significantly improved diagnosis and clinical practice in 68% of trials [7]. We also observed reasonable accuracy of diagnosis prediction for a very small but real medical data in our experimentation. The maximum benefit of CDSSs can be obtained in the future, when health care managements are fully electronic where real-time patient information will be available, as it simplifies the modifications of information in order to keep the system up to date. Developing a full-fledged and fully functional CDSS is quite hard, time-consuming, expensive, and laborious. Hence, we have started with a plug-in-based framework with minimal features with an intuition to add features in a timely manner. Adding new features, new and efficient algorithms, refining the decision-making inference engine, improving the knowledge-base construction process can be potential extensions. A better user interface reflecting the requirement of the clinicians with appropriate UI/UX survey and research can add greater value to the CDSS/EHR.

References

1. Begum S, Ahmed MU, Funk P, Xiong N, Folke M (2010) Case-based reasoning systems in the health sciences: a survey of recent trends and developments. IEEE Trans Syst Man, Cybernetics, Part C (Appl Rev) 41(4):421–434
2. Berner ES (2007) Clinical decision support systems. Springer, vol 233
3. Boyle T (2018) Medical transcriptions: medical transcription data scraped from mtsamples.com. https://www.kaggle.com/tboyle10/medicaltranscriptions
4. Jimison H, Gordon C, Berner E (2016) Clinical decision support systems: theory and practice. Health informatics, 3rd ed. Springer International Publishing, Cham, pp 163–79
5. Kawamoto K, Houlihan CA, Balas EA, Lobach DF (2005) Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. Bmj 330(7494):765
6. Kennebeck SS, Timm N, Farrell MK, Spooner SA (2012) Impact of electronic health record implementation on patient flow metrics in a pediatric emergency department. J Am Med Inf Assoc 19(3):443–447
7. Knoery CR, Heaton J, Polson R, Bond R, Iftikhar A, Rjoob K, McGilligan V, Peace A, Leslie SJ (2020) Systematic review of clinical decision support systems for prehospital acute coronary syndrome identification. Crit Pathways Cardiol 19(3):119
8. McGinn CA, Gagnon MP, Shaw N, Sicotte C, Mathieu L, Leduc Y, Grenier S, Duplantie J, Abdeljelil AB, Légaré F (2012) Users' perspectives of key factors to implementing electronic health records in canada: a delphi study. BMC Med Inf Dec Mak 12(1):1–13
9. Rothman B, Leonard JC, Vigoda MM (2012) Future of electronic health records: implications for decision support. Mount Sinai J Med: A J Trans Personalized Med 79(6):757–768
10. Rozenblum R, Jang Y, Zimlichman E, Salzberg C, Tamblyn M, Buckeridge D, Forster A, Bates DW, Tamblyn R (2011) A qualitative study of Canada's experience with the implementation of electronic health information technology. Cmaj 183(5):E281–E288
11. Wagholikar KB, Sundararajan V, Deshpande AW (2012) Modeling paradigms for medical diagnostic decision support: a survey and future directions. J Med Syst 36(5):3029–3049

The Impact of Collaborative Decision-Making in a Smart Manufacturing Environment: Case Study Using an Automated Water Bottling Plant

J. Coetzer, R. B. Kuriakose, H. J. Vermaak, and G. Nel

Abstract With the dawn of the Fourth Industrial Revolution (Industry 4.0), following three previous industrial revolutions, many disruptive technologies such as cloud computing, the Internet of Things (IoT), Internet of Services, Cyber Physical Systems (CPS) and big data are rapidly advancing throughout the world. Smart Manufacturing integrates these expeditious advancement of technologies related to the Fourth Industrial Revolution at the center of advancing automated structures. Preceding research within the Industry 4.0 research environment has for the most part focused on connecting machines and digital systems in an autonomous environment. However, the integration of humans within Industry 4.0 need to be considered as many industrial settings still make use of mixed environments. Within these environments the completion of processes are dependent on the cooperation between human workers and automated systems. At this time there is limited research on the development of collaborative decision-making where the human's adjustment and acknowledgment to the method is taken into consideration. A research gap identifying the lack of processes involving collaborative decision-making, has been established as the productive setup of an automated system relies greatly on the Human–Machine Interaction (HMI) between the machines and the human operator. The case study of a water bottling plant is utilized in this research to create a separation between the collection of tasks that need to be done by a machine as well as a human related to a Smart Manufacturing Environment. Even though the new technologies employed in Industry 4.0 may predict or detect a shift in procedure, human intervention and decision-making are still of critical importance. The paper initially discusses how the water bottling plant was modeled. Secondly, the paper then discusses collaborative

J. Coetzer (✉) · R. B. Kuriakose · H. J. Vermaak · G. Nel
Central University of Technology, Bloemfontein, Free State, South Africa
e-mail: jmarkram@cut.ac.za

R. B. Kuriakose
e-mail: rkuriako@cut.ac.za

H. J. Vermaak
e-mail: hvermaak@cut.ac.za

G. Nel
e-mail: guilnel@cut.ac.za

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_30

decision-making with reference to possible models for collaborative decision-making and introduces a HMI used for intervention in the automated process. The paper is rounded off by discussing the results and impact of collaborative decision-making in the automated procedure related to the case study.

Keywords Collaborative decision-making · Human–machine interaction · Smart manufacturing · Fourth industrial revolution

1 Introduction

Smart manufacturing is a widely used term to describe the current and future trends in manufacturing related to the Industry 4.0 environment [1]. According to the National Institute of Standards and Technology (NIST) Smart manufacturing is fully assimilated, collaborative manufacturing systems that respond in real-time to meet requirements and conditions in the factory, in the supply network and in customer demands [2]. In a Smart factory, machinery and equipment are able to improve processes through automation and self-optimization while functioning in a highly digitalized and connected environment [2].

Smart manufacturing integrates sensors, communication technology, computing platforms, data modeling, control and measurement by using rapidly developing technologies related to Industry 4.0 [3]. These technologies include Internet of Things (IoT), Cyber physical Systems (CPS), Internet of Everything (IoE), cloud computing, service-oriented computing, data science and Artificial Intelligence (AI). These technologies put the creating of automated systems and processes at the center of Industry 4.0 [4].

With the integrated processing and communication capabilities of the above mentioned technologies, a noteworthy concern is the result that it will have on the interaction between humans and machines [3]. Automated systems are not always exclusively autonomous; a degree of human interaction is practically constantly existent even though consideration is not always given to this challenge [3]. Efficient collaborative decision-making between the workforce and the automated system is imperative for the effective operation of an automated system [5].

Collaborative decision-making is a process whereby effective science-based decision-making is facilitated for the collaboration among designers, technical practitioners and users [6]. The human operator will still be involved in production or manufacturing processes whether it is for executing specific assembly tasks, supervising automated processes or intervening with processes when necessary [7]. Collaborative decision-making, in this context, thus refers to working together with a system or machine on a certain task to achieve a specific, common goal [8].

Previous research has extensively focused on the advancement and analyzing of new technologies associated with autonomous systems. In spite of this, there is a shortage of research done on collaborative decision-making, especially in mixed

environments, whereby the completion of a process is dependent on the collaboration between the human worker and the automated system [9].

Some of the studies that were performed in this field were carried out by Klumpp et al. [5] whereby a production logistics traffic control problem was introduced to create possible models for collaborative decision-making. Scenarios were executed to include firstly, human operators, secondly, automated robots and thirdly, a combination where collaboration between the human operator and the robot were introduced. These models served as actors in Industry 4.0 settings to scrutinize the critical role of collaborative decision-making and also human interaction in automated situations [5].

From a Smart manufacturing perspective, this research aims to implement collaborative decision-making within the manufacturing field. The explicit input of this paper is to investigate the impact of collaborative decision-making in a Smart manufacturing setting utilizing a case study of an automated water bottling plant. Two potential models will be introduced namely, machine only, and secondly, a collaboration between the human and the system. This is done by using the case study of a water bottling plant, henceforward referred to as the plant, which is modeled in Simulink and executed using Smart Manufacturing Units (SMU) [8].

This paper is structured such that first an overview is given of the existing water bottling plant and how it operates. Secondly, collaborative decision-making related to this study, is introduced and discussed. Thirdly, the results and findings are described and then rounded off with a conclusion.

2 Plant Overview and Modeling

This section gives an overview of the plant that is described in Sect. 1. The concept of the plant was aimed at producing bottled water for internal use at the Central University of Technology (CUT), Free State as, presented by the CUT management [10].

After an economic viability study was performed [8], the requirements for the plant was as follows:

- Able to bottle 5000 L of water a month
- Contain all minerals
- Should be able to source and store water
- Should be able to produce 500 and 750 ml bottles
- Fill and cap the 500 and 750 ml bottles

A three-dimensional Computer Aided Design (CAD) [11] model of the plant is portrayed in Fig. 1.

The different subsystems for the plant is show in Fig. 1, where (A) is the source and water storage tank, (B) is the bottle manufacturing and storage subsystem and (C) indicates the water filling unit.

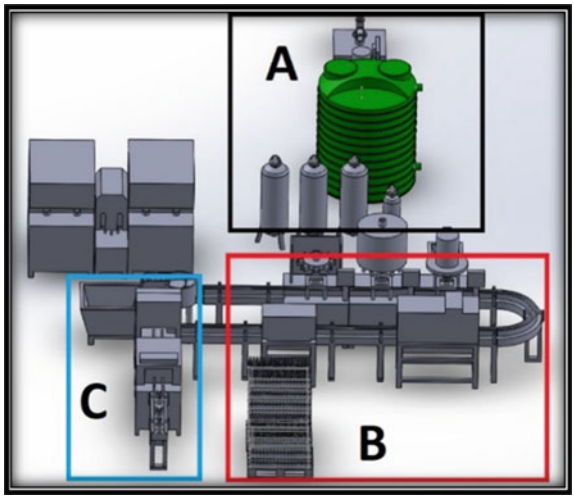


Fig. 1 A 3-D printed model illustrating the completed water bottling plant

The plant was modeled in Simulink and as part of the technical feasibility optimized using MATLAB. Input to the model is received via a web server through an interactive user interface that has been developed [8] where customers can place orders for either 500 ml or 750 ml bottles. The app requires customers to complete their personal details, indicate the desired bottle size and required date of delivery. The captured information is stored on a cloud server from where the MATLAB program can access and process the data to start filling the bottles. The user interface for the web application is shown in Fig. 2.

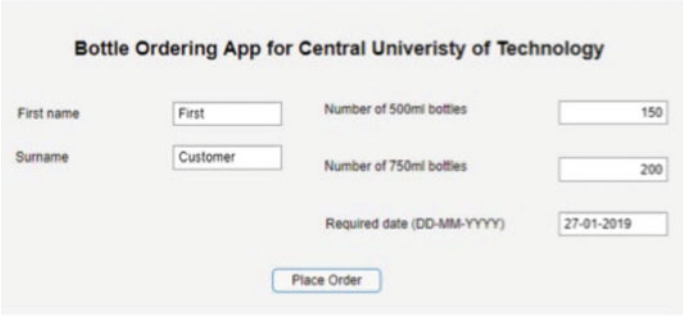


Fig. 2 The user interface for the online ordering of bottles

3 Collaborative Decision-Making in the Automated Water Bottling Plant

As described in the Introduction, the collaborative process of decision-making in the setting of this research, refers to humans working together with a system or machine while having a similar and specific mutual goal in mind. Potential models for human–robot interaction that examine the impact of collaborative decision-making were introduced by Klumpp, et al. [12].

For the purpose of Klumpp's research, a prototype for accident prevention in an autonomous traffic environment was introduced. According to the research, the prototype functioned satisfactory when the participants using the system were automated vehicles that followed well-defined rules precisely and could rapidly converse by means of using computer interfaces [5]. Initially the prototype was developed to be used only by automated systems but was expanded to incorporate humans and their interactions with the autonomous components. This led to proposing potential models for collaborative decision-making by means of human–machine interaction. The inclusion of humans that could operate vehicles remotely using a desktop application, was introduced for this purpose by adapting the control inputs of a vehicle to include collaborative decisions [13].

Three different categories were identified as potential models; human first, robot first and a hybrid where the human and robot collaborate. In the human first scenario, the vehicle controlled by the human was given priority which meant that the autonomous vehicle had to react swiftly to the inputs received from the human-controlled vehicle. The danger for potential collisions were high due to the short reaction times (for both robots and humans) when unforeseen actions occurred [5].

For the scenario where the robot was put in control, the vehicle controlled by the human was completely removed and the automated vehicle was forced to disregard the vehicle controlled by the human. As a result, the automated vehicles did not respond to the human-controlled vehicle which required the human to adjust their activities.

Thus, the task of avoiding collisions were the responsibility of the human. In this setup, collisions were the highest for human-controlled vehicles [5].

In the hybrid scenario, equal responsibility was given to the robot and the human for paying attention for avoidance of possible collisions. Although collisions still occurred, the impact of the hybrid or collaborative approach was that the number of collisions were the lowest between all three scenarios.

The aim of this paper is to test the impact of collaborative decision-making in a Smart manufacturing environment with the plant as the case study. As this is preliminary research, there will be two scenarios; machine first and a collaboration between the machine and humans. Both scenarios will be tested on the Smart Manufacturing Unit (SMU) described in Sect. 2.

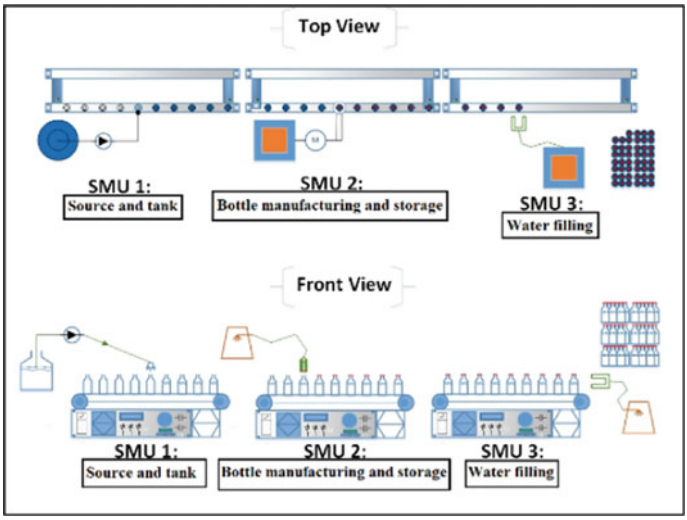


Fig. 3 The layout of the three SMU's of the water bottling plant

3.1 Machine Only

The Smart Manufacturing Unit is illustrated in Fig. 3 and displays the three SMU's used for the process of filling the water bottles. The first SMU consists of the source and tank, which contains water for filling the bottles, the second SMU manufactures and stores the bottles used for filling the water and the third SMU executes the filling of water into the bottles.

Both scenarios that will be tested as part of this research will be done on SMU 3. For the machine only scenario, the automated system is responsible for completing the task of filling either 500 ml or 750 ml bottles in an optimum time as orders are received via the web application. There are several constraints that are taken into account by the optimization model for the completion of the production process, which includes the number of orders obtained from the cloud, the water flow rate, the source tank water level and the number of 500 and 750 ml bottles in storage [14].

For the purpose of this research, only one constraint will be tested namely, the water level in the tank which should ideally not go below 25%. The SMU is programmed to automatically stop production when the water level hits 25% and will resume operation after the storage tank has reached full capacity.

3.2 Human–Machine Collaboration

In the human–machine collaboration scenario, there are different places where the human can fit into the production line. To understand this in detail, an exploded view

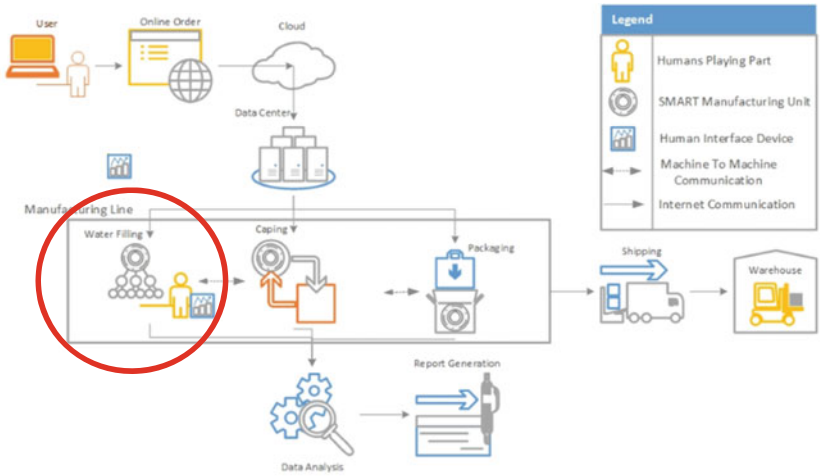


Fig. 4 The production line for the plant

of SMU 3 is provided in Fig. 4. As seen from Fig. 4, SMU 3 is split into the water filling, capping and packaging unit. A human intervention is possible in many areas such as at the placing of online orders, water filling, capping or packaging. For this research the human will intervene between the water filling station and the bottle capping station as depicted in Fig. 4.

With the intention of bringing the human into the production process, a HMI was developed and assigned for intervention by the human. The HMI is connected to the PLC and when the water reaches a specific level, an alert is created which requests whether the speed of production be slowed down or be sped up. At this point a decision can be taken by the human for adjusting the flow rate which is dependent on the water available in the source tank or to leave the speed unchanged. The hypothesis here is that the human-system collaboration will enable a quicker decision-making and therefore speed up production as opposed to a system only scenario.

4 Results and Discussions

The results of implementing an HMI to the existing plant which allows the human to intervene and take decisions that need to be made during the production process is discussed in this section. As mentioned in Sect. 3, the water bottling plant has been optimized to automatically fill either 500 or 750 ml bottles. The number of bottles that need filling is determined by customers placing orders online via the web application, thus the orders must be completed in the shortest time possible for timely delivery to customers.

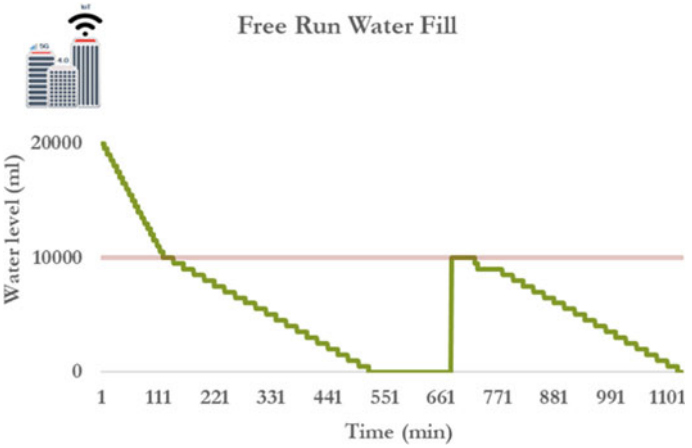


Fig. 5 Free run water fill indicating full system control

Several constraints are taken into consideration by the optimization model and as mentioned, the constraint tested for this research is the level of water available in the storage tank. A predetermined set of rules regulates the water in the storage tank with the condition that the water level must not get to 0% and is programmed to create an alert when the water level in the tank touches 25% [15].

A free run water fill, where there are no constraints attached to the process and the machine exclusively makes the decisions as depicted in Fig. 5, should not be considered as the level of the water in the storage tank should, ideally, not attain 0%. With the free run water fill, the bottles are being filled up until a 0% water level has been reached. As soon as this instance occurs, the whole filling process will come to a total standstill while the water level increases in order to restart the filling process. The free run scenario portrayed in Fig. 5 is showcased here to highlight the importance of having interventions either by the machine or by human intervention for optimization of the process.

The scenario depicted in Fig. 6 portrays the process of the system controlling the water level and taking all decisions for the filling process. When the water level reaches 25% or 10,000 ml, the process slows down, reaches a pause and waits for the water to be replenished in the storage tank. As soon as the water level has reached a specific level in the storage tank, the water filling process resumes. This operation is based on a predetermined algorithm. The total time to complete the order when the machine controls the water level, was approximately 1101 min, as portrayed in Fig. 6.

In the second scenario, the HMI is introduced to the plant. For the purpose of obtaining a holistic view of the manufacturing process for the plant, a Supervisory Control and Data Acquisition System (SCADA) is implemented. A SCADA is utilized for the regulation and observing various aspects of manufacturing plants environments [16]. A SCADA system operates by working with indicators that uses

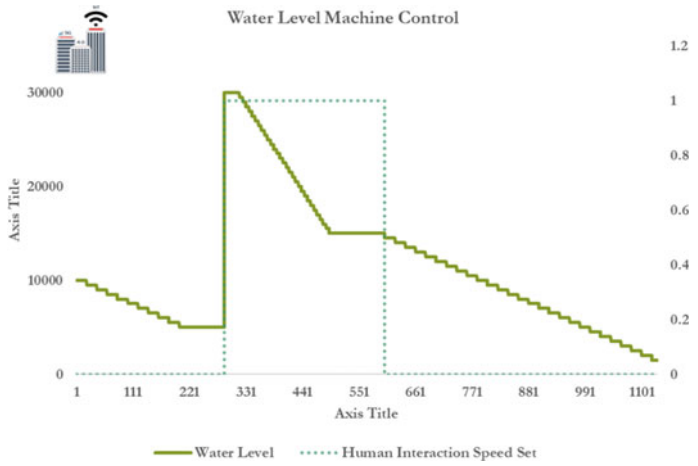


Fig. 6 Water level as controlled by system only

communication channels for providing the user with remote controlling of equipment for a specific structure [17]. In comparison, HMI's provide an efficient means for communicating with hardware and can thus be considered as a subsection of a SCADA system.

The SCADA will provide the human with a complete picture of the variants pertaining to the plant such as the water available in the storage tank, the flow rate of the water and the amount of bottles available. Based on these values, the human will be able to take informed decisions via the HMI.

During the HMI controlled operation, two scenarios will be tested; one where the water flow to the storage tank will be delayed and the process needs to be slowed down, and another scenario where the water flow to the storage tank is imminent, hence the process can be sped up.

Figure 7 indicates the process where water filling is slowed down. The process starts at 20,000 ml and continues filling bottles until the water level reaches 10,000 ml or 25% as shown in Fig. 7. At this point the human anticipates that the water is not enough to complete the required orders in time and hence intervenes via the HMI by overriding the process to slow down the filling of bottles until the water level of the storage tank reaches a certain set level. The speed is constantly being slowed down as it waits for the water level to rise although the bottles are still being filled, only at a slower pace. As depicted, the time to complete the order took 881 min.

The scenario where the speed is increased, depicted in Fig. 8, is in the long run the ultimate scenario where the water level is dropping and reaches a position where the system is programmed to pause and wait for an increase in the water level. At this point the human is allowed to intervene by means of the HMI to indicate that the machine ought to increase the flow rate and continue with the filling process as the operator knows in advance that water will shortly reach the system. The machine

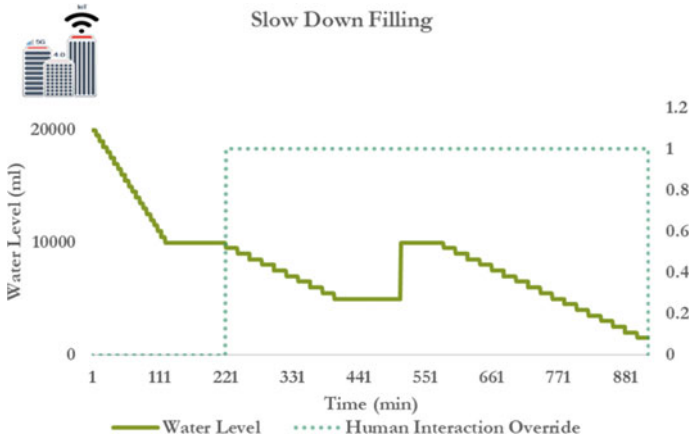


Fig. 7 Human interaction overrides to slow down filling

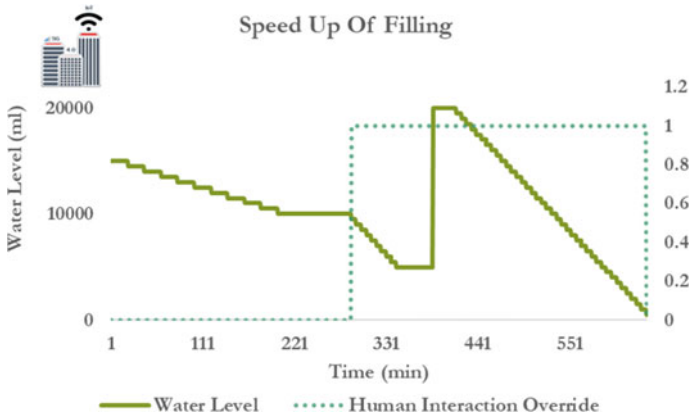


Fig. 8 Human interaction overrides to speed up filling

does thus not have to slow down and the human interaction overrides to speed up the process.

The entire process of filling the bottles and completing orders, take a certain time. When the machine controls the process and makes the decisions, as depicted in Fig. 6, the process of filling the bottles take slightly longer than in the other two scenarios where human interaction occurs for overriding the process. For the filling of bottles using only machine control, the process took approximately 1101 min, in comparison with the scenarios where human intervention took place.

When human intervention takes place via the HMI, the two scenarios described involves the human that needs to make a decision for indicating whether the process should slow down or speed up. Even though the human intervention indicates that the process should slow down, as depicted in Fig. 7, the time to complete the process

is less than that of where the machine makes the decision to wait for the water level to increase before completing the process, which took approximately 881 min. The same is true where the human intervenes when it is known that water is immediately available and the process should speed up, as indicated in Fig. 8. For this scenario the process took about 551 min.

5 Conclusion

This paper introduces collaborative decision-making by means of a HMI to a Smart manufacturing environment utilizing the case study of an automated water bottling plant. The aim of the research was to examine whether collaborative decision-making between the machine and the human will have a positive impact on the optimization of the manufacture process.

The paper initially gives an overview of the plant. As illustrated in Fig. 1, the plant was modeled in Simulink and optimized using MATLAB. The process of ordering the product via a web application, developed for customer orders, was part of the overview and briefly explained. The graphical user interface for the web application is depicted in Fig. 2.

The next section of the paper discusses collaborative decision-making with reference to potential models for collaborative decision-making as introduced by a comparable research project. For this research, two scenarios were presented namely the machine only and secondly, a collaboration between the machine and the human. A HMI was presented to the production process at the SMU for filling the water bottles of the in order to bring the human into the process as depicted in Fig. 4. The scenarios were implemented and yielded significant results as described in Sect. 4.

The results, showcased in Figs. 7 and 8, show that the process was at least 20% faster with human intervention in comparison with the machine only scenario. This is in correlation with Klumpp's [5] experiment where the collaboration between the robot and human resulted in less traffic collisions than that of the robot only scenario. It can thus be deduced that in this particular case, collaboration between the machine and the human utilizing an HMI has a positive impact on the production process.

The purpose of this research is to see where we can bring meaningful contribution by humans to a Smart manufacturing environment. As a result, humans will not be placed in all aspects of the Smart manufacturing plant. For this experiment the human intervention utilizing the HMI took place at the point between the water filling station and the capping station. The fact that the HMI was tested on only one SMU unit of the plant is a limitation of this study.

It is important to note that the inclusion of humans in the Smart manufacturing environment creates opportunities for workers who may lose their jobs due to Industry 4.0 technologies to be re-skilled to be part of a collaborative decision-making environment. More research will need to be done on the kind and level of re-skilling that will be required to ensure the worker can be re-deployed to a different job within the Industry 4.0 environment.

The research is thus limited to firstly identifying areas where human contributions would hold value during the production process and warrants further research to arrive at optimized processes in an automated environment.

References

1. Kusiak A (2018) Smart manufacturing. Int J Prod Res 56(1–2):508–517. https://doi.org/10.1080/00207543.2017.1351644
2. Helu et al (2016) Enabling smart manufacturing technologies for decision-making support. In: ASME design engineering technical conference
3. Coetzer J, Kuriakose RB, Vermaak HJ (2020) Collaborative decision making for human technology interaction—a case study using an automated water bottling plant. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/1577/1/012024
4. Nagy J, Oláh J, Erdei E, Máté D, Popp J (2018) The role and impact of industry 4.0 and the internet of things on the business strategy of the value chain-the case of hungary. Sustain 10(10). https://doi.org/10.3390/su10103491
5. Klumpp M, Hesenius M, Meyer O, Ruiner C, Gruhn V (2019) Production logistics and human-computer interaction—state-of-the-art, challenges and requirements for the future. Int J Adv Manuf Technol. https://doi.org/10.1007/s00170-019-03785-0
6. Ulibarri N (2018) Collaborative model development increases trust in and use of scientific information in environmental decision-making. Environ Sci Policy 82(February):136–142. https://doi.org/10.1016/j.envsci.2018.01.022
7. Faber M, Bützler J, Schlick CM (2015) Human-robot cooperation in future production systems: analysis of requirements for designing an ergonomic work system. Procedia Manuf 3(AFHE):510–517. https://doi.org/10.1016/j.promfg.2015.07.215
8. Kuriakose RB, Vermaak HJ (2019) Customized mixed model stochastic assembly line modelling using Simulink. Int J Simul Syst Sci Technol 2–7. https://doi.org/10.5013/IJSSST.a.20.S1.06
9. Coetzer J, Kuriakose R, Vermaak H (2021) Devising a novel means of introducing collaborative decision-making to an automated water bottling plant to study the impact of positive drift. In: Fong S, Zhuhai M, Joshi A, Dey N (eds) Lecture notes in networks and systems 154ICT analysis and applications, vol 154. Springer Nature Singapore Pvt Ltd., Singapore, 2020, pp 661–669
10. 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 Crypt 23(2):329–336
11. Kuriakose RB, Vermaak HJ (2019) Optimization of a customized mixed model assembly using MATLAB/Simulink. J Phys Conf Ser 1201(1)
12. Klumpp M (2018) Automation and artificial intelligence in business logistics systems: human reactions and collaboration requirements. Int J Logist Res Appl 21(3):224–242
13. Coetzer J, Kuriakose RB, Vermaak HJ (2020) Collaborative decision-making for human-technology interaction-a case study using an automated water bottling plant. J Phys Conf Ser 1577(1). https://doi.org/10.1088/1742-6596/1577/1/012024
14. Kuriakose_2019_J._Phys.__Conf._Ser._1201_012017.pdf
15. Automation P (2018) Differences between SCADA and HMI. [Online]. Available. https://info.premierautomation.com/blog/differences-between-scada-and-hmi
16. Krambeck D (2015) An introduction to SCADA systems. [Online]. Available. https://www.allaboutcircuits.com/technical-articles/an-introduction-to-scada-systems/

Analysis of Social Assistance During the COVID-19 Pandemic

Ahmad Ahmad, Achmad Nurmandi, Isnaini Muallidin, and Mohammad Jafar Loilatu

Abstract This research focuses on the Ministry of Social Affairs in distributing social assistance. This study aims to find out how the process of distributing social aid carried out by the Ministry of Social Affairs has been going well following the procedures. And during the distribution, there were any obstacles that hindered the distribution process so that the distribution of social assistance was not right on target. The method used in this research is descriptive qualitative using online news media as data processed through NVivo 12 Plus software. This approach was chosen to see the extent to which the Ministry of Social Affairs was distributing social assistance to people affected by the COVID-19 pandemic with indicators of resources, bureaucratic structure, communication, and commitment as comparison material. Based on the data analysis that has been carried out, the bureaucratic structure is a major problem that causes delays in the distribution of social assistance with a percentage of 81.25%, not without reasons for lack of accountability and the occurrence of abuse of office by bureaucrats who control the distribution of social assistance and use the situation to commit acts of corruption. Furthermore, Resources became the second factor that caused the distribution of social assistance not to go well with a presentation of around 37.50%. Furthermore, the results of data analysis with the commitment indicator, which is not much different from the bureaucratic structure, become the third reason for the problems in the distribution of social assistance due to factors from the stakeholders in the distribution of social service who lack a strong commitment with a value presentation of around 21.25% as the primary basis in implementing responsibility. In contrast to the results of data analysis from the communication indicator, 18.75% is the lowest problem in the distribution of social assistance by the Ministry of Social Affairs to the results of data analysis from the communication indicator of 18.75%.

Keywords Social assistance · Policy · COVID-19 · Communication

A. Ahmad (✉) · A. Nurmandi · I. Muallidin · M. J. Loilatu
Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: enreahmad4@gmail.com

A. Nurmandi
e-mail: nurmandi_achmad@umy.ac.id

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_31

1 Introduction

With the outbreak initially starting in Wuhan, China quickly spread worldwide, including Indonesia. The corona 19 virus emerged and became one of the emergencies. According to Hua and Shaw [1], coronavirus 19 is a humanitarian emergency, where this outbreak appeared in Wuhan, China in 2019 at the beginning of December, and the news began to spread at the end of December, or at the beginning of January 2020, in the third week of January—detected as an emergency. On January 31, 2020, WHO finally declared this a Public Health Emergency of International Concern (PHEIC) and announced a pandemic on March 11, 2020, as of March 24, 2020. Gullslett et al. [2] stated that the coronavirus 19 pandemic was a health emergency in international history. That in international history the Corona Virus 19 pandemic is an unprecedented public health emergency.

The public does not know when it will come, mainly because the spread is speedy. It has become a trending topic worldwide in the last two years, namely the coronavirus or better known as coronavirus disease 2019 COVID-19. Gude and Muire [3] explaining that the spread of the corona 19 virus can be through direct contact with exposed people and possibly with tiny droplets of saliva produced when coughing, sneezing, or talking. This coronavirus is a complex disease with indications of attacking the respiratory tract, and the rate of transmission is very fast through the air or the atmosphere [4].

Meanwhile, as stated by Astuti [5], this virus is indicated to have originated from live animals traded in a traditional market in Wuhan, China. Severe acute respiratory syndrome can attack lung cells due to many stored receptor entries, namely angiotensin converting enzyme. The presence of this virus in the host cell initiates various protective responses leading to pneumonia and acute respiratory distress syndrome.

The state must anticipate a shield that should be installed from the start so that the long-term economy does not become a new problem for the state. The more a country does not follow the rules of the health protocol to maintain a distance, the greater the spread of the COVID-19 virus. Quarantine plays a significant role in controlling the coronavirus. Patients who have been exposed or diagnosed with the coronavirus should be placed in separate places until the recovery process [6]. The government's and community's responses to preventive measures, such as school closures, working from home, especially formal sector workers, delays, and cancelations of various government and private events, have slowed the wheels of the economy [7].

These challenges undeniably will be pretty impactful for Indonesia. The economic impact can be even worse for a long time. This means that the more a country enforces protocol rules for its people to implement community activity restrictions (PPKM) will impact the community's economy, especially if this continues in the long term, thus, maintaining economic stability and community survival. The government has helped by making various forms of policy. Gelb and Mukherjee [8] not only in Indonesia but in other countries too have launched an unprecedented aid package to protect the economy in their respective countries.

An efficient bureaucracy is essential in the twenty-first century because the government and public administration play an indispensable role in facilitating economic and social change and the growing awareness that government must act more swiftly [9]. In Indonesia itself, many types of assistance have been distributed to people affected by the corona 19 pandemic in the form of cash, necessities, and relief in the form of cutting electricity tariff bills in the hope of reducing the burden on the community.

Most of the heads of households have lost their jobs and are vulnerable to receiving social benefits. Therefore, they will likely find it difficult to rely on immediate assistance from the community [10]. The service provided by the government aims to ensure the availability of basic needs and provide social protection, especially for several groups affected by COVID-19.

2 Literature Review

2.1 Public Policy Implementation

George Edward III emphasized that the main problem of public administration is the lack of attention to the issue of policy implementation [11]. According to Edward, without effective policy implementation, policymakers' decisions will not succeed. For policy implementation to be effective, Edward [11] suggests four main issues that must be considered, namely communication (communication), resources (resources), commitment (disposition or attitude), and bureaucratic structure.

Communication (communication) is related to how policies are communicated to the organization or the public, the availability of policy resources, attitudes, and responses from the parties involved, and the organizational structure of implementing policies. Resources (resources) relate to the availability of supporting resources, especially human resources. A vital resource aspect, in this case, is the skill of policy implementers who will implement the policy effectively. Commitment (disposition) relates to the willingness and dedication of the implementers to implement the policy effectively. The bureaucratic structure is related to the suitability of the bureaucratic organization that implements public procedures. In this case, what needs to be maintained is how to prevent bureaucratic fragmentation in its implementation because such a structure will hinder the performance of the public policy.

3 Methodology

The method used in this research is descriptive qualitative research. Online media is a sample of data sources taken through N Capture for NVivo 12 Plus with Web Chrome which is processed using NVivo 12 Plus software. The captured data containing

Table 1 Online media sources

No	Media	Source link	Whole
1	CNN	https://www.cnnindonesia.com	5
2	Detik.com	https://www.detik.com	15
3	Kompas	https://www.kompas.com	15
4	Media Indonesia	https://mediaindonesia.com	5

Source Compiled by the author, 2021

information from an online media account are coded at this stage. The data are processed with the crosstab feature to automatically calculate the necessary main statistical tests with significant comparisons of indicators. The feature used in this process is the node feature which aims to classify the data in the study.

Files related to research that have been imported are then read to find sentences resulting from the analysis to be grouped according to node indicators. Furthermore, the most frequently used words or terms can be revealed by using the word cloud feature. This study examines how the Ministry of Social Affairs distributes social assistance during the pandemic through online news media by measuring indicators of resources, bureaucratic structure, communication, and commitment (Table 1).

4 Discussion

In carrying out life in the current situation, of course, there are many challenges that the community must go through to survive. Without reason, the many problems that occurred in our beloved country, starting from the unfinished spread of the coronavirus and the emergence of disasters that rocked one after another, make people's lives even more alarming. To survive, of course, people cannot rely on themselves but need government assistance as we know that the government is one form of organization that works and carries out the task of managing the government system and setting policies in maintaining the sovereignty of the state and ensuring the welfare of its people. The Ministry of Social Affairs of the Republic of Indonesia is responsible for ensuring success in distributing social assistance to communities affected by the COVID-19 pandemic commitment.

4.1 Bureaucratic Structure

In essence, the bureaucracy is present as the implementation of the government system, both in responding to various problems and in providing services to the

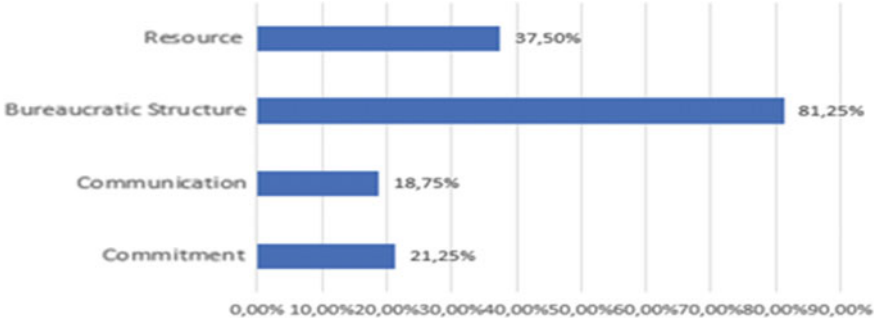


Fig. 1 Matrix coding, NVivo 12 plus indicators of implementing social assistance distribution policies. *Source* Compiled by the author, 2021

community. The importance of one of the conditions of a professional bureaucracy is to provide services to the community (public service). Still, in reality, it is different, where public trust is broken by irresponsible elements who have taken advantage of their position.

Based on the results of the data in Fig. 1, it can be seen that the indicator of bureaucratic structure is the most dominating problem that results in policy implementation in the distribution of social assistance not going well, not without reason that the problem arises because of small things that are considered trivial such as bureaucrats are too busy with internal affairs so that matters regarding services to the community are neglected, lack of honesty, and accountability so that they take advantage of the situation to commit corrupt acts.

4.2 Resource

The results of the data analysis in Fig. 1 show that the indicator of resources is in second place with a value presentation of 37.50%, which means that it is the biggest problem after the bureaucratic structure that hinders the distribution of social assistance. Resources are an essential element in running the government, with the current state of society still needing attention from the government, both at the center and the local level. Therefore, people who are considered to have high quality and credibility are expected to carry out their duties and responsibilities in serving, assisting, and improving people's lives through policies and programs implemented by the government. However, the facts show that there are still socio-economic gaps in the strata of people's lives, one of which is the unequal distribution of social assistance carried out by the Ministry of Social Affairs.

On the other hand, two factors hinder the distribution of social assistance on resource indicators, namely geographical conditions and managers' inability. Figure 2 shows the data on the failure of managers, in this case, human resources, to

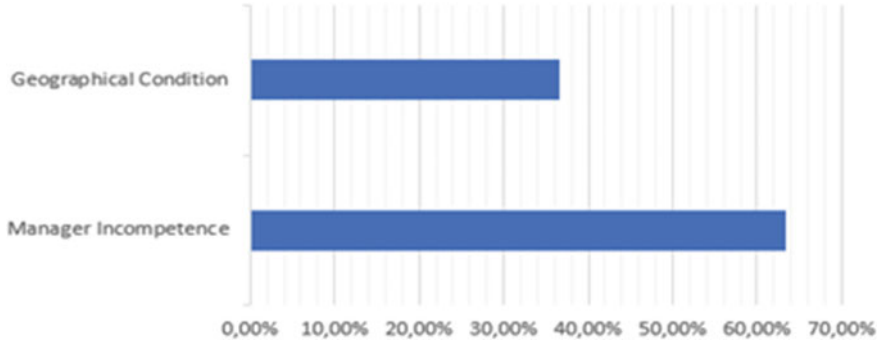


Fig. 2 Causes of weak resources. *Source* Compiled by the author, 2021

carry out their duties properly. These problems are very closely related to geographical conditions where certain areas are difficult to access by distributors in the direction, including the interior of Papua, the interior of NTT, and many other regions. The lack of access to the interior is an advantage for the individuals involved in distributing social assistance, manipulating data by using the names of social service recipients with other families and even providing social assistance to their close relatives. With actions like this, the distribution of social assistance is not channeled on target.

4.3 Commitment

Within the scope of government, the terms servant and service user are not something new to hear. In line with that, commitment is the most crucial element in service delivery. Therefore, the community will assess the role of commitment in implementing public services as service users. The commitment to implementing public services should be prioritized, at least including professionalism and ethics in providing services to the community. Therefore, as best as possible, the standard operating procedure (SOP) in a service provider agency, without a commitment, it will not work as it should.

The results of data analysis in Fig. 1, with a percentage value of 21.25%, indicate that the commitment indicator is the third most significant problem that causes the distribution of social assistance to be constrained. The distribution of social service did not go well due to a lack of honesty and the absence of a strong commitment by the managers. Commitment is needed in carrying out the duties and responsibilities to control oneself to stay on the right path. One of the shreds of evidence of the lack of commitment in the authorities who manage social assistance is the evidence of corruption in social service involving the minister of social affairs.

4.4 Communication

In development or service in government institutions, communication is one of the essential factors that are the main trigger for this success to be achieved. Communication is a liaison tool between humans in the organization's scope. However, communication is very closely related to government. It becomes the flow of delivery and acceptance of the market that links each other between the government as the sender and the community as the recipient of the message. Thus, to measure the success of the distribution of social assistance, communication is one of the four indicators that causes the distribution of social aid not to go well. In Fig. 1, we can see that communication is in fourth place with a score of 15.75%. In terms of numbers, of course, it is not a considerable value among other indicators. Still, it has a significant impact on the loss of people who need this assistance.

In measuring how much influence this communication has, it is the cause of the unsuccessful distribution of social assistance during the COVID-19 pandemic. Furthermore, the authors analyze the communication indicators with derivatives, namely the sender and recipient of the message, as in Fig. 3.

From the results of data analysis in Fig. 3, the comparison is quite far, namely 75% and 27%, which means that the government as the messenger is less in providing information related to the distribution of social assistance to the community as the recipient of the message. This problem can be seen from the lack of socialization regarding the terms and conditions that must be met by prospective recipients of social assistance, giving rise to pros and cons in the community. In addition, people's assumptions about the system of favoritism and kinship are very clearly shown so that the distribution of social assistance is not evenly distributed.

In line with that, the cause of problems in the distribution of social assistance lies not only in areas that are difficult to reach, lack of communication, or lack of socialization. Figure 4 explains that the complex, irrelevant social assistance data,

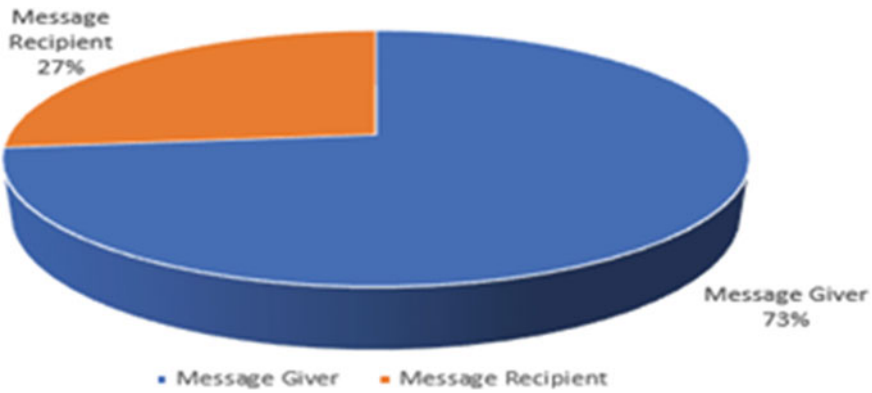


Fig. 3 Influence of communication on the distribution of social assistance. *Source* Compiled by the author, 2021

Fig. 4 Word frequency, implementation of social assistance distribution. *Source* Compiled by the author, 2021

and the discrepancy between central government data and blood government data that are difficult to track are the primary triggers for this social assistance problem to begin. This can be proven based on the results of word frequency analysis which shows that the word social assistance data have a presentation that often appears.

Data inaccuracies have always been a classic problem in every distribution of social assistance. It is not just a question of inaccurate social assistance recipient data. Several factors are suspected to be the cause of the data inaccuracy. The social welfare integrated data (STKS) have not been connected to the population identification number (NIK) and the delay and weakness in the data verification process. In addition, the number of social assistance programs launched by the Ministry of Social Affairs during the COVID-19 pandemic resulted in the overlapping of these programs, resulting in the emergence of new vulnerable groups of people who were not previously registered in the list of recipients of social assistance, in the end, they did not receive any social service.

5 Conclusion

Based on the results of data analysis, it can be concluded that the Government of Indonesia's policy in distributing social assistance to assist and reduce the burden on the affected community during the COVID-19 pandemic is still considered to be less than satisfactory for the community because there are problems involving elements of the community. An irresponsible person as the holder of authority, in this case, the Ministry of Social Affairs who intentionally commits acts of abuse of power and

takes advantage of the situation to commit acts of corruption which results in the distribution of social assistance not going well, and the distribution is not right on target.

In realizing the social assistance distribution program, several factors trigger problems in the distribution of social assistance. First, there is a lot of complex data but a lack of resources to improve data at both the central and regional levels. Second, the distribution system of social assistance is not transparent, and the supervision is considered weak so that problems are found, such as the identity of the population of the recipients of social aid being invalid. From some of these problems, there has been unrest among the people who feel that social assistance is not evenly distributed, starting in terms of time, target recipients, distribution areas, and unclear procedures and requirements for receiving social service.

References

1. Hua J, Shaw R (2020) Corona virus (Covid-19)'infodemic' and emerging issues through a data lens: the case of china. J Environ Res Public Health. https://doi.org/10.3390/ijerph17072309
2. Gullslett MK, Kristiansen E, Nilsen ER (2021) Therapists' experience of video consultation in specialized mental health services during the COVID-19 pandemic: qualitative interview study. JMIR Hum Factors 8(3). https://doi.org/10.2196/23150
3. Gude VG, Muire PJ (2021) Preparing for outbreaks—implications for resilient water utility operations and services. Sustain Cities Soc 64. https://doi.org/10.1016/j.scs.2020.102558
4. Zaharah Z, Kirilova G, Windarti A (2020) Impact of corona virus outbreak towards teaching and learning activities in Indonesia. Jurnal Sosial dan Budaya Syar-i [Online]. Available. http://journal.uinjkt.ac.id/index.php/salam/article/view/15104. Accessed 11 Oct 2021
5. Astuti I (2020) Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): an overview of viral structure and host response. Diab Metabol Syndr Clin Res Rev [Online]. Available. https://www.sciencedirect.com/science/article/pii/S1871402120300849. Accessed 11 Oct 2021
6. Said C, Batchelor F, Duque G (2020) Physical activity and exercise for older people during and after the coronavirus disease 2019 pandemic: a path to recovery. J Am Med Directors Assoc [Online]. Available. www.ncbi.nlm.nih.gov/pmc/articles/PMC7269946/. Accessed 11 Oct 2021
7. Muzakki F (2020) The global political economy impact of Covid-19 and the implication to Indonesia. J Soc Polit Sci 1(2) [Online]. Available. http://e-journal.unas.ac.id/index.php/jsps/article/view/11. Accessed 11 Oct 2021
8. Gelb A, Mukherjee A (2020) Digital technology in social assistance transfers for COVID-19 relief: lessons from selected cases. CGD policy paper [Online]. Available. https://www.ictworks.org/wp-content/uploads/2020/09/lessons-learned-digital-technology-social-assistance-programs.pdf. Accessed 11 Oct 2021
9. Yuniza ME, Rebecca AG (2021) Red tape phenomenon of social securities distribution during Covid-19: a socio–legal analysis. IKAT Indonesian J SE Asian Stud 4(2):113–127. https://doi.org/10.22146/ikat.v4i2.58137
10. Kartseva MA, Kuznetsova PO (2020) The economic consequences of the coronavirus pandemic: which groups will suffer more in terms of loss of employment and income? Popul Econ. https://doi.org/10.3897/popecon.4.e53194
11. Affrian R (2012) Kebijakan Publik by Eko Handoyo (z-lib.org), Semarang: Widya Karya, p 323

Toward an Agile and Transformational Government, Through the Development of the Tangerang LIVE Application (Case Study of Tangerang City, Indonesia)

Ahmad Syukri, Achmad Nurmandi, Isnaini Muallidin, Danang Kurniawan, and Mohammad Jafar Loilatu

Abstract This research aims to see how far the application of "Tangerang LIVE" is in realizing the principles of agile governance in the Tangerang City Government. The current era of technology is a challenge for every government to have information and communication technology (ICT) capabilities in the application of public services. The Tangerang City Government answered this challenge by making innovation in an application called "Tangerang LIVE". The method in this study is a descriptive analysis that explains a clear picture of the application of agile governance principles to applications made by the Tangerang City Government. Researchers analyzed the data using the "Tangerang LIVE" application found on the google play store for users' smartphone Android. This study's data collection techniques were carried out through secondary data searches obtained from books, articles, news, comments, government publications (applications), and journals. This research is necessary because it can see how far the Tangerang City Government is in realizing the principles of agile governance through public services. The results of this study indicate that of the three principles, agile governance discussed, the Tangerang City Government has implemented several principles of agile governance, namely based on quick wins and simple design and continuous refinement. However, on the focus of the systematic and adaptive approach, it is still not implemented optimally. This can be seen through the number of user complaints related to the verification process, which takes a long time to respond after registering on the application.

A. Syukri (✉) · A. Nurmandi · I. Muallidin · D. Kurniawan · M. J. Loilatu
Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: ahmad2016.jbi@gmail.com

A. Nurmandi
e-mail: nurmandi_achmad@umy.ac.id

I. Muallidin
e-mail: isnainimuallidin@gmail.com

D. Kurniawan
e-mail: kurniawand949@gmail.com

M. J. Loilatu
e-mail: jafar.loilatu@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_32

Keywords Application · Agile governance · Tangerang LIVE · E-government

1 Introduction

Entering the era of information technology transformation, a government with information and communication technology (ICT) skills is needed to convey messages to the public. ICT uses digital and places a new government bureaucracy called e-government [1]. Services in the digital form delivered online are an effort to increase efficiency for the community in accessing information [2]. The application of e-government in government using various online applications will boost public participation [3]. Thus, the ease of obtaining information is expected to create good governance and foster trust in the community.

The government is encouraged to implement organizational management agile to realize good governance [4]. Obtaining information easily and quickly becomes a challenge for the government in serving the community's needs. The Tangerang City Government carries out the application of ICT by making an application called "Tangerang LIVE". This is one form of innovation carried out by the Tangerang City Government in realizing agile governance. Application as software functions to perform various forms of work regarding the application, use, and addition of data [5]. Through the online-based application media "Tangerang LIVE" which the city government launched, it is expected to provide all the information needed by the people of Tangerang City.

"Tangerang LIVE" is an application portal created by the Tangerang City Government to support various services needed by the community. Until now, the "Tangerang LIVE" application has been continuously developed since it was first launched in August 2016 [6]. There are various features provided, such as LAKSA (Your Suggestion Box Aspiration Service), Single Number Emergency Call Service 112, Online Licensing Service, and other parts that can make it easier for the people of Tangerang City. The use of digital-based applications in the public service sector is an effort to improve performance and provide involvement to the community to realize the value of transparency [1]. In addition, government innovations carried out by making the "Tangerang LIVE" application aim to create a government agile.

Agile governance emerged as a new, broad, and multidisciplinary paradigm that focuses on the performance of an organization and is related to competitiveness, which until now has not been explored in-depth [7]. So that in the current technological era, ICT skills from the government are needed to create a responsive and efficient public service [8]. So, this study aims to find out how the Tangerang City Government applies principles of Agile Governance through the "Tangerang LIVE" application as a form of service to the community.

2 Literature Review

2.1 E-Government

Gartner defines e-government as a continuous effort in optimizing a service. Organizations transform a service using technology, the Internet, and new media [9]. E-government can be interpreted as a form or model of a government system that is based on the use of digital technology. Such as administrative work, public services, financial management, and so on will be controlled in one system [10]. The e-government application is used to provide better service from the government. Such as more efficient public services, emphasizing cost control, and taking advantage of the advancements that ICT continues to offer [11].

According to Indrajit, e-government is a new interaction mechanism between the government and the community or organizations interested. The process carried out involves using information technology such as Internet media, which aims to improve the quality of public service [10]. The implementation of e-government carried out in government circles aims to make it easier in the service process to the community. In addition, the government, the community, and the private sector (companies) can collaborate to realize effectiveness and efficiency in service.

2.2 Agile Government

Agility is an organization's capacity that is efficient and effective in using human resources, both internally and externally, to create activity values with higher results [12]. An agile organization must manage and consider the strategies needed and desired. The concept of agility is also getting attention in the field of government. Because the competitive environment is becoming more unpredictable, and the competition is getting tougher. This leads to applying values in agile software development in the process of good governance [13]. Agile governance is the ability of human resources to feel, adapt, and respond quickly and sustainably to change in their surroundings. Applying agile capabilities to governance aimed at delivering faster, better, and more affordable values to the service processes implemented by an organization [14]. A principle is essential in agile governance. Luna et al. suggest that there are six principles of agile governance: (1) good enough governance, (2) business-driven, (3) human focused, (4) based on quick wins, (5) systematic and adaptive approach, (6) simple design and continuous refinement [15].

This study uses 3 (three) principles of agile governance proposed by Luna et al. [14], including: (1) based on quick wins, a success that is achieved quickly must be motivated to achieve further success. (2) Systematic and adaptive approach, a team needs to expand its intrinsic capabilities to respond to changes quickly and systematically. (3) Simple design and continuous refinement, the team must have the ability to create simple designs and provide fast and constantly improving results.

3 Research Method

This research was conducted to see the importance of agile governance principles in the Tangerang City Government in providing services to the community through the "Tangerang LIVE" application. This research is qualitative research with a descriptive analysis approach that explains a clear picture of the application of agile governance principles in applications made by the Tangerang City Government. Researchers analyzed the data using the "Tangerang LIVE" application found on the google Play Store for smartphone Android. The data collection technique in this study is in accordance with that proposed by Silalahi which was carried out through searching secondary data obtained from books, articles, news, comments, government publications (applications), and journals [15]. This study focuses on the use of the "Tangerang LIVE" application in realizing the principles of agile governance in the City of Tangerang, Indonesia.

4 Result and Discussion

4.1 Overview of the Tangerang LIVE Application

The "Tangerang LIVE" application is one of the innovations made by the Tangerang City Government. Through this application, it is hoped that it will make it easier for people to receive information. In addition, the "Tangerang LIVE" application also offers various application features that can help people from different aspects. Such as matters regarding online licensing, emergency calls, information about COVID-19, and many other uses that can be utilized. Making applications is an application of e-government that is expected to provide many benefits to the public. Misuraca suggests that the implementation of e-government brings many advantages, one of which is creating a new environment for the community that can answer all problems quickly and accurately and can face various changes globally [16].

The emergence of the digital revolution has made many government agencies seek to utilize the Internet to realize ease in implementing a public service [17]. Digital applications are a tangible manifestation of implementing e-government carried out by the Tangerang City Government. The "Tangerang LIVE" application was first launched in August 2016. This application can be used on smartphones and downloaded through the Play Store for Android users. The development of the "Tangerang LIVE" application is also being carried out to improve the quality of usability in various features to impact the community significantly. The following is the initial screen when opening the "Tangerang LIVE" application (Fig. 1).

Based on the appearance on the home page, you can see various features in the "Tangerang LIVE" application. Some of the components contained in the application require the public to register to be used, especially for the people of Tangerang City. In addition, the appearance of the start page on the application looks attractive and is

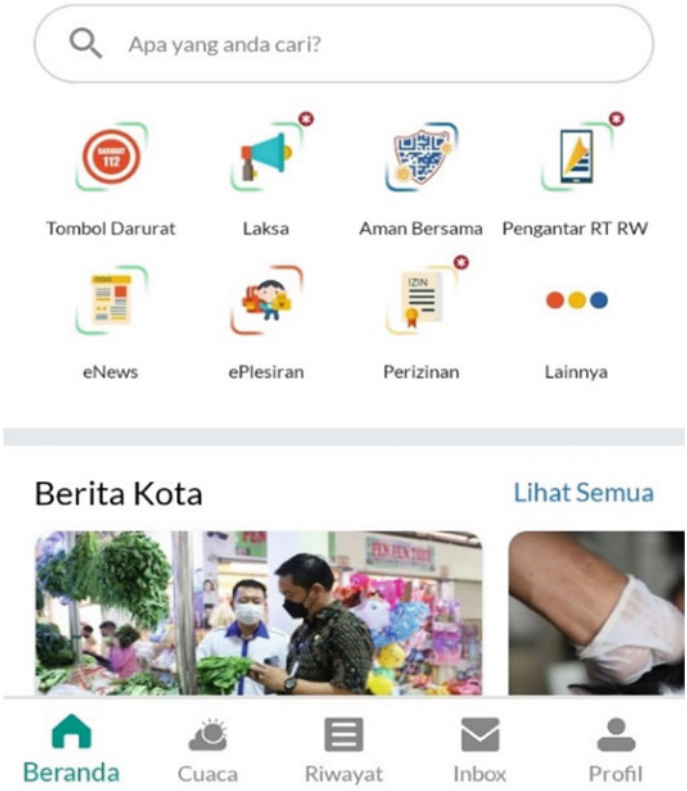


Fig. 1 Display on the homepage of the "Tangerang LIVE" application [18]

also easy to observe. Furthermore, on the start page of the application, it also places daily news at the bottom to provide information to the public regarding updates regarding the latest news in the City of Tangerang.

Since its launch in August 2016, the "Tangerang LIVE" application has attracted much public interest in its use. Until 2021, there is an increase in the download of the "Tangerang LIVE" application (Fig. 2).

Based on the data above, through the Tangerang City Communication and Information Office (Diskominfo), in March 2021, there was an increase in the download of the "Tangerang LIVE" application. The number of downloaders in January was 525,197 and increased in February to 540,960, then raised in March, which reached 546,111 people [19]. Making the application "Tangerang LIVE" represents a goal in implementing the concept of e-government and realizing Agile Governance in the Tangerang City government. Therefore, to determine the extent to which the Tangerang City Government applies agile governance, the researchers analyzed the three principles of [14], which are as follows.

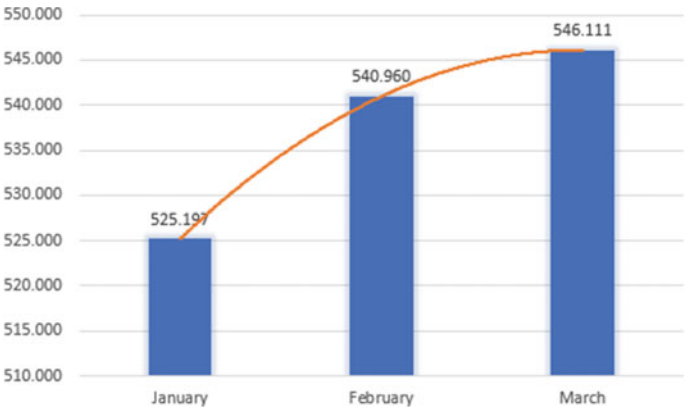


Fig. 2 Increase in the number of Tangerang LIVE application downloaders in 2021. *Source* http://katakota.com

4.2 Based on Quick Wins

According to Luna et al. [14], explaining the principle based on quick wins that a success that is achieved quickly must be used as motivation as an impetus to achieve further success. The Tangerang City Government continues to develop the "Tangerang LIVE" application. Various new features are made to serve the community quickly and practically. The era of the COVID-19 pandemic has made multiple sectors in the government sector experience complex problems. Among them are problems in the health and economic fields. The Tangerang City Government, through the "Tangerang LIVE" application, added several new features to provide easy access for the community (Fig. 3).

There are several features added to the "Tangerang LIVE" application. Such as in the field of health services by adding news features about Covid-19 and Covid-19 vaccine locations in the Tangerang City area. In addition, there is the Tangerang Portal feature in the economic sector, which aims to meet shopping needs with a cod (cash on delivery) payment process. This means that the addition of the new feature is a manifestation of the implementation of the Tangerang City Government on principle based on quick wins.

4.3 Systematic and Adaptive Approach

The second principle regarding a systematic and adaptive approach is that the Tangerang City Government, through the "Tangerang LIVE" application, must respond quickly and systematically to application updates. There are several complaints from the people of Tangerang City, which can be seen in the comments column on the Android Play Store (Fig. 4).

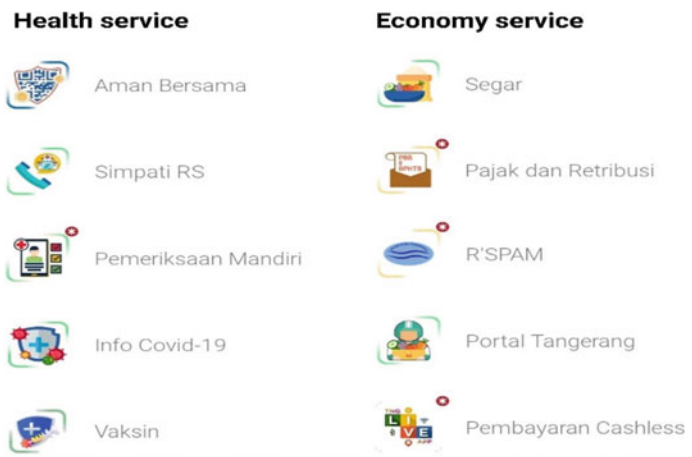


Fig. 3 Adding application features to health and economic services [18]

Fig. 4 Public complaints on registration problems [18]

Several complaints often appear in the comments column on the Play Store for users of Android. Among them are complaints about the verification process, which takes quite a long time to respond to when registering. This means that this application is considered not optimal based on a systematic and adaptive approach. The government is expected to react quickly to create good accessibility when used by the people of Tangerang City.

4.4 Simple Design and Continuous Refinement

The third principle is simple design and continuous refinement. The "Tangerang LIVE" application must have a simple design and provide fast and continually improving results (Fig. 5).

Based on the above data related to COVID-19 information that can be seen in appearance, the "Tangerang LIVE" application is considered simple and easy to

Fig. 5 Display of COVID-19 information in Tangerang City [18]

understand by the public. A non-monotonous design can provide comfort for application users. In addition, the COVID-19 information on the "Tangerang LIVE" application provides the latest data presentation and continues to be updated every day. Therefore, the application is considered capable of helping the people of Tangerang City, which can be measured through application downloader data which continues to increase in 2021.

5 Conclusion

Making the "Tangerang LIVE" application is an innovation from the Tangerang City Government to realize a government agile. This research shows that the "Tangerang LIVE" application already embodies several principles of agile governance. First, based on quick wins, namely the Tangerang City Government continues to add new features by looking at the latest problems felt by the community. Additions in the form of health service features and also economical services. Second, the principle of simple design and continuous refinement is to apply a simple and attractive appearance to applications and information data updated every day, such as information on COVID-19 in Tangerang City. However, there are still shortcomings that need to be improved on the application, as in the third principle, namely the systematic and adaptive approach. There are many user complaints regarding the verification process, which takes a long time to respond after registering on the application. Therefore, the Tangerang City Government is expected to respond quickly to complaints from application users to create good accessibility when used by the people of Tangerang City.

References

1. Adinda PP, Suzianti A (2018) Redesign of user interface for E-government application using usability testing method. In: ACM international conference proceeding series, pp 145–149. https://doi.org/10.1145/3290420.3290433
2. Curtis S (2019) Digital transformation—the silver bullet to public service improvement? Public Money Manag 39(5):322–324. https://doi.org/10.1080/09540962.2019.1611233
3. Razak MRR, Ahmad J, Mustanir A, Madani M, Idhan A (2021) Effects of ICT application (information communication and technology) towards the transformation of regional bureaucracy. IOP Conf Ser Earth Environ Sci 717(1). https://doi.org/10.1088/1755-1315/717/1/012003
4. VernandaR (2020) Kesiapan Indonesia menuju agile governance. Konf Nas Ilmu Adm 4.0 (KNIA 4.0) 1–6 [Online]. Available. http://180.250.247.102/conference/index.php/knia/article/view/147.
5. Hartati S, Kristiana Dewi NA, Puastuti D, Muslihudin M, Setio Budi N (2017) Sistem Aplikasi EDUCHAT STMIK PRINGSEWU Berbasis ANDROID Sebagai Media Komunikasi dan Informasi. J Nas Teknol dan Sist Inf 3(1):143–152. https://doi.org/10.25077/teknosi.v3i1.2017.143-152
6. Chairul Fikri FH (2018) Aplikasi Tangerang Live 4.0 Mudahkan Pelayanan Masyarakat. Jakarta, 16 Apr 2018
7. Luna AJ, Marinho ML, de Moura HP (2020) Agile governance theory: operationalization. Innov Syst Softw Eng 16(1):3–44. https://doi.org/10.1007/s11334-019-00345-3
8. Mutiarin D, Moner Y, Nurmandi A (2019) The adoption of information and communication technologies in human resource management in the era of public governance. Public Policy Adm 18(2):346–362. https://doi.org/10.13165/VPA-19-18-2-12
9. Jayashree S, Marthandan G (2010) Government to E-government to E-society. J Appl Sci 10(19):2205–2210. https://doi.org/10.3923/jas.2010.2205.2210
10. Muhammad M, Djumlani A, Paselle E (2019) Studi Tentang Pemanfaatan E-Procurement Pada Dinas Pekerjaan Umum Kabupaten Kutai Timur. J Paradig 8(1):22. https://doi.org/10.30872/jp.v8i1.2924

11. Vidmar R, Kotzé E, Merwe MV, Mentz J (2019) An approach to design and evaluation of e-Government theory for South Africa
12. Teece DJ, Peteratd M, Leih S (2016) Dynamic capabilities and organizational agility. Calif Manage Rev 58(4):13–35
13. Ylinen M (2021) Incorporating agile practices in public sector IT management: a nudge toward adaptive governance. Inf Polity 26(3):251–271. https://doi.org/10.3233/IP-200269
14. Luna AJ, Kruchten P, de Moura HP (2015) Agile governance theory: conceptual development. pp 1–22 [Online]. Available. http://arxiv.org/abs/1505.06701
15. Halim FR, Astuti F, Umam K (2021) Implementasi Prinsip Agile Governance Melalui Aplikasi PIKOBAR di Provinsi Jawa Barat. Kolaborasi J Adm Publik 7(1):48–67
16. Irawan B (2013) Studi Analisis Konsep E-Government: Sebuah Paradigma Baru dalam Pelayanan Publik Bambang. J Paradig 2:174–201
17. Abdel-Basset M, Zhou Y, Mohamed M, Chang V (2018) A group decision making framework based on neutrosophic VIKOR approach for e-government website evaluation. J Intell Fuzzy Syst 34(6):4213–4224. https://doi.org/10.3233/JIFS-171952
18. Pemerintah Kota Tangerang (2016) Tangerang LIVE. https://play.google.com/store/apps/details?id=id.go.tangerangkota.tangeranglive&hl=in&gl=US
19. Katakota.com (2021) Pengguna Aplikasi Tangerang LIVE Semakin Meningkat. Tangerang, 06 Apr 2021

The Successful Use of the PeduliLindungi Application in Handling COVID-19 (Indonesian Case Study)

Akhdiva Elfi Istiqoh, Achmad Nurmandi, Isnaini Muallidin, Mohammad Jafar Loilatu, and Danang Kurniawan

Abstract This study aims to determine the success of using the PeduliLindungi application in Indonesia. The government created the PeduliLindungi application in dealing with COVID-19 in Indonesia, integrated nationally. The PeduliLindungi application is one of the government's breakthroughs in handling the COVID-19 pandemic. The PeduliLindungi application has several challenges: how people trust the application, the role of the application in daily life, and how the government can make a strategy so that people are sure of using the PeduliLindungi application. The data used in this study is a qualitative descriptive data and Twitter data related to public responses. This study used a QDAS approach to analyze the NVIVO Plus 12 software. The results showed that the PeduliLindungi application provided convenience to users through the features on the application offered. The PeduliLindungi application components include COVID-19 test results, Ehac, teledokter, COVID-19 statistical information, and vaccination registration. The government also has a good strategy in convincing the public to use the PeduliLindungi application by making rules that explain the obligation to use the PeduliLindungi application in daily life during the COVID-19 pandemic. This rule makes the number of PeduliLindungi application users increase each month. However, the PeduliLindungi application also received criticism from application users regarding complaints in using the PeduliLindungi application, such as application errors, data leaks, and user data errors. This criticism can be used as an evaluation material for the government to improve the quality of the PeduliLindungi application.

Keywords Applications · COVID-19 · PeduliLindungi

1 Introduction

For the first time, the COVID-19 virus appeared in December 2019 and quickly spread worldwide. Even more than 121 countries, including Indonesia, were affected. The

A. E. Istiqoh (✉) · A. Nurmandi · I. Muallidin · M. J. Loilatu · D. Kurniawan
Department of Government Affairs and Administration, Jusuf Kalla School Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: akhdivaelfiistiqoh@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_33

virus killed more than 4000 people in China at that time. In 2020, the World Health Organization (WHO) announced that the coronavirus (COVID-19) had become a global pandemic [1]. In recent years, subjective well-being among mobile app users has attracted researchers' interest because of its principal role in improving daily life, especially during the current coronavirus (COVID-19) pandemic [2]. Several countries have shown interest in digital contact tracing applications since mid-March 2020. In Europe, there has been widespread media coverage of academic and corporate developments and government agencies' statements and actions regarding their interest [3].

Some other countries are also doing the same thing, developing an application similar to Singapore, India, Greece, Australia, and China. To record and do tracing to the community, a group of researchers from Southern Illinois University (SIU) recently built an information representation device that uses GPS data to show users the location of known COVID-19 cases. Google and Apple have also teamed up to develop contacts and applications powered by Bluetooth technology. These methods have effectively collected extensive and accurate data [4]. Singapore developed an application called Trace Together to identify people who may have been exposed to COVID-19. Trace Together uses Bluetooth signals between users to track [5].

Indonesia has created an application called PeduliLindungi, released by the Ministry of Communication and Information (Kominfo) in April 2020 to assist the government in handling the COVID-19 pandemic in Indonesia. Therefore, the Ministry of Communication and Information (Kominfo) recommends every Indonesian people to download the PeduliLindungi application [1]. Monitoring the spread of COVID-19 in Indonesia, which the Government of Indonesia handles in various ways, is through the android application installed by the Google Play Store. An application made by the government, which is a forum for handling and community participation to share location data while traveling so that tracing contact history with COVID-19 sufferers in Indonesia can be carried out immediately. This application is certainly a discussion that is often discussed by the public with their individual opinions and opinions through the Cares Protect application on Google Play [6]. Based on the background explanation above, the author wants to see the success of the PeduliLindungi application in handling COVID-19 in Indonesia.

2 Literature Review

2.1 Application Development Success Strategy

In northern Nigeria, where food insecurity is so severe, they illustrate how applications can provide real time, high frequency, and accurate, validated data on the evolution of commodity prices [7]. As a result of the current COVID-19 pandemic, recovering patients may have PTSD. Alex Poulos explained in his research the importance of using an application that can help these former COVID-19 patients increase

their emotional resilience and their ability to recover from trauma after being exposed to COVID-19 [8]. To succeed today, the government needs a strategy in its implementation. An excellent digital strategy is a strategy that can provide progress and evaluation as required [9].

2.2 Digital Transition in Handling COVID-19

Social restrictions led to a severe economic downturn and a massive spike in digitization. The required reduction in face-to-face contact dramatically increases the use of digital media in the economy and society [10]. During a pandemic, digital technology can reduce or even solve many challenges, thereby improving healthcare delivery. Work is currently underway to address acute needs that are a direct or indirect consequence of the pandemic (e.g., applications for patient tracking, remote triage emergency services, etc.) [11]. Technology can change the way many industrial sectors and processes operate [12]. Indicators that can support the digital transition procedure are (1) the value of the trust, (2) the transformative role of digital, and (3) the identification of digital strategies [13]. The idea is that digitization is more of a human process than a technical one. A common analogy is that 20% of digitization involves technology, and 80% involves people, with the people component essentially being the change management process [14].

2.3 Community Participation in Implementing the Digital Transition

Public trust in health care broadly defined, trust in the healthcare system has to do with how the public perceives the design and the actors within it related to their ability to provide services and seek the best interests of their clients. Trust is important because it impacts various health behaviors, including adherence, and ultimately affects the power of the healthcare system to meet its goals [15]. Digital participation is expected to increase the legitimacy and quality of decisions, both concluded during the process and based on input from participatory initiatives [16]. Digital involvement from the community can assist in the evaluation of system development [17].

3 Method

This study uses a QDAS approach to analyze the NVIVO Plus 12 software. The data used are descriptive qualitative method data, namely application data that explain the usefulness of the features in the PeduliLindungi application and public responses

related to the application seen from Twitter data. This study focuses on finding out how successful the PeduliLindungi application is in handling the COVID-19 pandemic in Indonesia.

4 Results and Discussions

4.1 Value of Trust in PeduliLindungi Applications

The digital transition in Indonesia during the COVID-19 pandemic began with creating an application called PeduliLindungi by the Indonesian Government to stop and reduce the spread of COVID-19 cases in Indonesia. The use of the PeduliLindungi application has been structured so that all Indonesian people can use it. Before this application can be helpful and used by the community, the government must convince the public to want to take advantage of the application that has been provided because trust from the community will be the first benchmark whether this PeduliLindungi application will be efficient and achieve the targets that have been set in its manufacture.

It can be seen from the data above that the PeduliLindungi application is based on the number of users. The graph has consistently increased in the last 12 months in the previous few months. 4–5 million period users (July–December 2020) 32.8–55 million user's period (July–September 2021), and there is still the possibility to continue to grow. This graph proves that the value of public trust in the PeduliLindungi application is increasing rapidly. These data have shown that the PeduliLindungi application has won the public's trust. The public already knows that the PeduliLindungi application is vital during the COVID-19 pandemic.

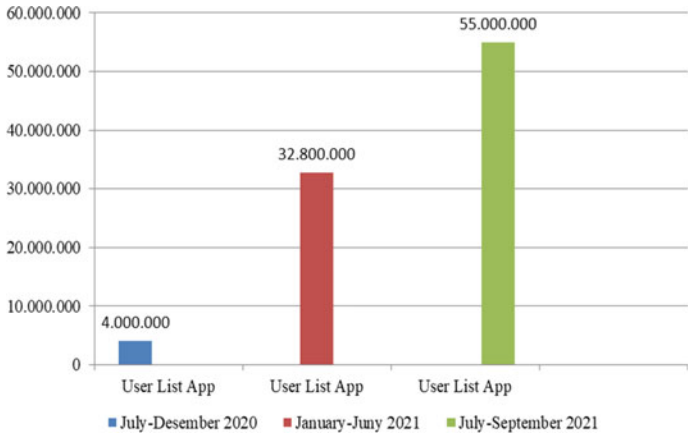


Fig. 1 Number of users of PeduliLindungi applications

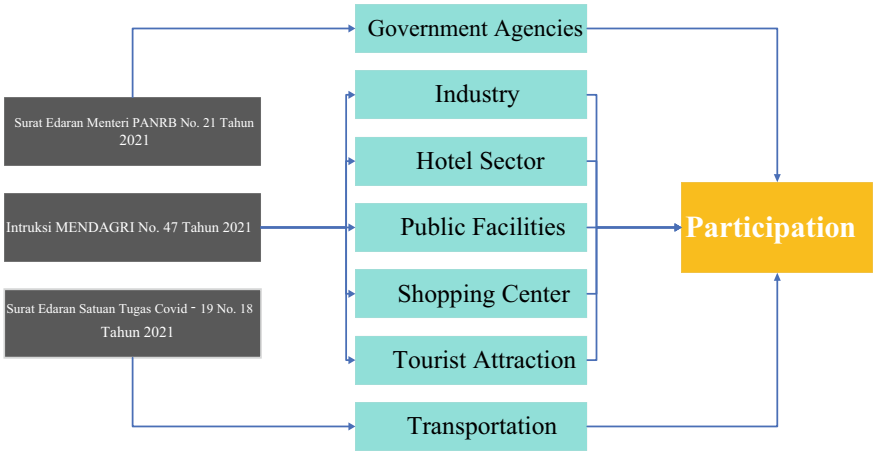


Fig. 2 Government regulations governing the use of PeduliLindungi applications

This user upgrade is inseparable from the digital strategy implemented by the government, such as requiring the use of the PeduliLindungi application in transportation such as airplanes. The government has also started trying to implement the PeduliLindungi application as a condition for entering crowded centers or public places such as malls, tourist attractions, and others. On the other hand, the government also issues vaccine certificates in the PeduliLindungi application to download this application. Through the policies made by the government, the Indonesian people can have the desire to use the PeduliLindungi application.

It can be seen from the data above that the government has made several policies or regulations which discuss the use of the PeduliLindungi application. The minister's circular for utilizing state apparatus and bureaucratic reform number 21 of 2021 requires the PeduliLindungi application in government agencies throughout Indonesia. The Minister of Home Affairs Instruction number 47 of 2021 requires the PeduliLindungi application in many public sectors, like an industrial area, hotels, public facilities, shopping centers, and tourist attractions. Government policies issued in the circular letter of the COVID-19 task force number 18 of 2021 also require the application of Care to Protect in the fields of land, sea, and air transportation. The government requires the use of the PeduliLindungi application due to the need for massive community tracking to assist the government in handling the COVID-19 pandemic in Indonesia.

4.2 Health Services Through PeduliLindungi

Contact tracing apps are advanced software tools that can help control the spread of COVID-19. The PeduliLindungi application has become an alternative by the

Table 1 Comparison before and after the application of PeduliLindungi

Before	After
There is no information about the zoning map of the spread of COVID-19	There is information about the zoning map of the distance of COVID-19
Health checkup must go directly to the health center	Health checkup can be done independently from home
Vaccine registration must come to vaccine location	Vaccine registration can be done from home
Vaccine certificate still in paper form	Can download vaccine certificate on the app
Airworthy permit must show two different letters and must always carry it everywhere	Flight-worthy permission has been directly connected to the application user's cellphone

community in daily life mobility; because the PeduliLindungi application has tracking and reduction features when the user is confirmed to be COVID-19, the implementation of the application provides many benefits for the community during this pandemic. The reduction in face-to-face services recommended by the government will also be carried out well because of the government's digital transition efforts.

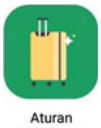
It can be seen from several comparisons before and after the PeduliLindungi application that the researchers presented. It was found that the digital transition during the pandemic, in this case, the PeduliLindungi application in Indonesia, provided many benefits to the Indonesian people in the mobility of daily life during the pandemic. For example, information on the zoning map of the spread of COVID-19 in Indonesia can make people care when traveling to an area. Then, health checks that can be done from home, registration of vaccines that can be done without having to come directly to the vaccine location, digital passports can make airworthy permits flown by the government to be more effective and efficient.

The PeduliLindungi application has many features that can help people's mobility in daily life and features that allow the government to track the spread of COVID-19 cases in Indonesia. The components in the PeduliLindungi application are developed from time to time. Previously, it was only to download vaccine certificates, and now, new features have been developed that will be useful in the community. These additional features are excellent because they help the community get optimal health services, especially amid the coronavirus pandemic. As is known, all face-to-face activities, especially in places with large crowds, should be avoided as much as possible during the pandemic to prevent the spread of COVID-19.

4.3 Community Response as a Form of Digital Participation

The digital transition has provided easy access for the public during the COVID-19 pandemic, one of which is in the form of application-based services offered by the government, which play an essential role in preventing the expansion of the spread

Table 2 PeduliLindungi application features and functions

Feature	Function
	This feature is used when application users enter public facilities such as shopping centers, hotels, and banks. This feature is helpful for the manager of public facilities to regulate the visitor density. This feature is also used to check the completeness of travel documents
	This feature is integrated with national vaccine data, and application users use this feature to download vaccine certificates for users who have been vaccinated
	This feature is integrated with data from the COVID-19 examination test results. This feature is used to view the results of COVID-19 tests that application users have carried out
	The Ehac feature is a form of digitizing health documents. This feature makes it easier for users who travel by air. They don't have to bother with vaccination checks and health tests because all the requirements have been integrated into digital documents in the PeduliLindungi application
	This feature serves to view the application user's travel history, whether it is a trip using transportation and places you have visited
	This feature serves to conduct health consultations with health workers remotely without directly coming to a health facility. This feature is also helpful in conducting independent health checks for application users
	This feature is used to find the nearest health facility and COVID-19 test site, polymerase chain reaction (PCR) or ANTIGEN. This feature is helpful because application users can see information about the nearest health facility in their vicinity

(continued)

COVID-19 outbreak, especially in the health sector. The presence of the PeduliLindungi application must gain the community's trust in its application. Public trust can be seen in the people who use the application during this pandemic. Community participation can be seen from the community's response to the application used or the public's response. Many benefits that the presence of the PeduliLindungi application has provided. The author finds there are also some unfavorable responses from the community regarding the technical side of the PeduliLindungi application.

Table 2 (continued)

Feature	Function
	This feature is used to see the number of COVID-19 cases around application users. By turning on location information, statistics will appear regarding the number of recovered patients, isolated patients, and patients who died in the vicinity of where the application user lives
	This feature is used to register vaccinations. Application users do not have to come directly to the vaccination location to vaccinate, register at home through the PeduliLindungi application, and wait for the notification of the vaccine schedule and location for PeduliLindungi application users

It can be seen in Fig. 3, which shows that the public response to the use of the PeduliLindungi application that most often appears is the application error reaching 52%, then data error getting 28%, and data leakage going 20%. It can be concluded that these three things are the findings by researchers in the application of the PeduliLindungi application in Indonesia and become a conversation in the form of responses by application users.

Figure 4 shows that the most discussed public responses are data errors, applications errors, and data leaks. This data error is related to data from users, which often results in errors such as vaccine certificate data that do not come out, then incorrect data on prospective vaccine recipients. This error problem is related to applications that often have errors when users use the PeduliLindungi application when entering public facilities such as stations and others. The application often has errors that

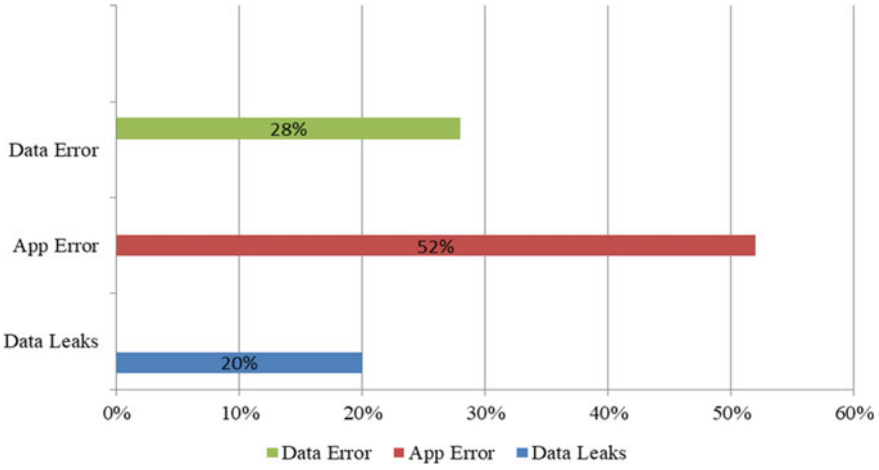


Fig. 3 Technical error people talk about

Fig. 4 Public response to PeduliLindungi applications

impact users, such as missing the train and the buildup of queues because of the application error. Then, several times, the application cannot be accessed because of the error.

Then, this data leak arose due to public anxiety about the security of their data in the form of a national population identification number and other data that users entered when registering on the application. This insecurity arises due to the leaking of the Indonesian president's national population identification number to the public so that concerns occur from people who use the PeduliLindungi application. All forms of public response regarding the PeduliLindungi application are a form of community participation in maximizing the government's digital transition process. This response can also be used as a reference by the government to further improve the features in the application and fix existing errors and deficiencies.

5 Conclusion

The use of the PeduliLindungi application during the COVID-19 pandemic in Indonesia was considered successful. The government's efforts to create the PeduliLindungi application have succeeded in helping people's mobility in daily life and assisting the government in handling the COVID-19 pandemic. This success can be seen from the increase in the number of application users every month, the positive impact provided through the features and functions of the application, as well as the success of the government in convincing the public to use the PeduliLindungi application. Meanwhile, in its implementation, it turns out that complaints are still found in the form of public responses from PeduliLindungi application users, such as application errors, data leaks, and user data errors. This criticism can be used as evaluation material for the government to further improve the quality of

the PeduliLindungi application in the future. This research still has several shortcomings, starting from the source of application user data that cannot be obtained in the PeduliLindungi application and limitations in collecting respondents to see the failures of the PeduliLindungi application. The researcher hopes that there will be further research to discuss the weaknesses of the PeduliLindungi application and develop further as an evaluation material for the government in improving the quality of the PeduliLindungi application.

References

1. Kurniawati KM, Riana D, Arfian A, Rahmawati E, Heriyanto (2020) Public acceptance of PeduliLindungi application in the acceleration of corona virus (Covid-19) handling. In: 1st International Conference on Advanced Information Scientific Development, ICAISD 2020, vol 1641, no 1
2. Aboelmaged M, Ali I, Hashem G (2021) Mobile apps use for wellness and fitness and university students' subjective wellbeing. Inf Dev
3. Lucivero F, Hallowell N, Johnson S, Prainsack B, Samuel G, Sharon T (2020) COVID-19 and contact tracing apps: ethical challenges for a social experiment on a global scale. J Bioeth Inc
4. Mohanty S, Harun AI Rashid M, Mridul M, Mohanty C, Swayamsiddha S. (2020) Application of artificial intelligence in COVID-19 drug repurposing. Diabetes Metab Syndr Clin Res Rev 14(5):1027–1031
5. Lai SHS, Tang CQY, Kurup A, Thevendran G (2021) The experience of contact tracing in Singapore in the control of COVID-19: highlighting the use of digital technology. Int Orthop 45(1):65–69
6. Mustopa A, Hermanto A, Pratama EB, Hendini A, Risdiansyah D (2020) Analysis of user reviews for the PeduliLindungi application on google play using the support vector machine and naive Bayes algorithm based on particle swarm optimization. In: 2020 5th international conference on informatics and computing (ICIC 2020)
7. Adewopo JB, Solano-Hermosilla G, Colen L, Micale F (2021) Using crowd-sourced data for real-time monitoring of food prices during the COVID-19 pandemic: insights from a pilot project in northern Nigeria. Glob Food Sec 29
8. Reyes AT (2020) A mindfulness mobile app for traumatized COVID-19 healthcare workers and recovered patients: a response to 'the use of digital applications and COVID-19.' Community Ment Health J 56(7):1204–1205
9. Ross JW, Sebastian IM, Beath CM (2017) Winter 2017 issue how to develop a great digital strategy. MIT Sloan Manage Rev 58(2):2017
10. Dannenberg P, Fuchs M, Riedler T, Wiedemann C (2020) Digital transition by COVID-19 pandemic? The German food online retail. Tijdschr voor Econ en Soc Geogr 111(3):543–560
11. Golinelli D, Boetto E, Carullo G, Nuzzolese AG, Landini MP, Fantini MP (2020) How the COVID-19 pandemic is favoring the adoption of digital technologies in healthcare: a literature review. J Med Internet Res 0–2
12. Della Valle F, Oliver M (2020) Blockchain enablers for supply chains: how to boost implementation in industry. IEEE Access 8:209699–209716
13. Della Valle F, Oliver M (2021) A guidance for blockchain-based digital transition in supply chains. Appl Sci 11(14)
14. Randall L, Berlina A (2019) Governing the digital transition in Nordic regions: the human element, vol 4
15. Peters FFY, Dexnell (2016) Public trust in health care. Int J Health Plan Manage 31(2)

16. Hofmann S et al (2021) Electronic participation to cite this version: HAL Id: Hal-03201953. Springer International Publication
17. Literat I, Kligler-Vilenchik N, Brough M, Blum-Ross A (2018) Analyzing youth digital participation: aims, actors, contexts, and intensities. Inf Soc 34(4):261–273

Application of the JKN Mobile Application in Improving the Quality of Health Services During the COVID-19

Randa Gustiawan, Achmad Nurmandi, Isnaini Muallidin, and Mohammad Jafar Loilatu

Abstract This study aims to determine the application of the JKN Mobile Application in improving the quality of health services during the COVID-19 period. In this study, researchers used the E-GovQual indicator. Researchers in conducting research using a qualitative descriptive approach. This research data was taken from the JKN Mobile Application, Twitter, Online Media, JKN Website, and Play Store and then presented descriptively. This study indicates that the JKN Mobile Application has improved the quality of public services in the health sector during the COVID-19 period. There are several assessment indicators. First, the Ease of use indicator, namely, the JKN Mobile Application, has made it easier for the public to access the presence of several features, and this application can overcome queue buildup and prevent the spread of COVID-19. Second, the indicator of trust, namely, the JKN Mobile Application, has won the public's trust. Third, reliability, namely, the JKN Mobile Application, which has provided easy access for the community to obtain health services. The JKN Mobile Application also provides features that support the community's needs in health services. Fourth, Content and Information Display, the features provided get a positive response from the community because it is easy to understand; people need to choose the health services they need. Finally, the Citizen Support, JKN Mobile Application provides information and complaints features that are useful for serving the problems and difficulties of the community in accessing health services and information needed.

Keywords E-government · E-service · Mobile JKN · COVID-19

R. Gustiawan (✉) · A. Nurmandi · I. Muallidin · M. J. Loilatu
Department of Government Affairs and Administration, Jusuf Kalla School of Government, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: randagustiawan17@gmail.com

A. Nurmandi
e-mail: nurmandi_achmad@umy.ac.id

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_34

1 Introduction

Corona virus disease or COVID-19 has become a global pandemic after being announced by the World Health Organization on March 11, 2020, the pandemic first appeared in Wuhan City, China [1]. Since it first appeared many changes have occurred in the lives of the world's people, including in Indonesia. The government announced the first case of COVID-19 virus infection in Indonesia on March 2, 2020 [2]. The pandemic has a significant impact in health services [3, 4]. In addition, health services are focused burdened with preventing transmission and handling cases of COVID-19, plus the implementation of the "lockdown" system [5].

The COVID-19 pandemic has given important lessons in terms of bureaucratic governance. In this COVID-19 era, the bureaucracy is obliged to maximize the use of technology, information and communication in all matters in the government sector [6]. ICT has an important role in all areas of human life, including technological solutions, and efficient applications in e-service processes [7]. In the health service sector, the use of information technology is considered very important to avoid the accumulation of people who need health services and prevent the expansion of the spread of COVID-19.

The transformation of digital services experienced a significant spike marked by face-to-face to electronic services in the government sector called "*Electronic Government*". During the COVID-19 pandemic, services will be increasingly difficult to reach with the existence of social distancing policies, physical distancing and the implementation of the lockdown system in various regions which causes a new order in the bureaucracy to provide services and the community to get services. This condition demands that service delivery must be optimized with technology applications. This condition is in front of us for a change toward a new normal life [6]. The presence of Mobile JKN provides new innovations to improve services quality to the community, especially in the health sector [8, 9]. Through the mobile application, people can get services anytime and anywhere without having to come directly to the office [10]. So that it can overcome the accumulation of queues and prevent the expansion of the spread of COVID-19.

In this digitalization era and in the midst of a pandemic where almost all aspects of life are creating, using and utilizing digital media in the government sector called E-Government. Likewise, health services experience technological adoption and innovation which will certainly speed up the service process. In this study, researchers will analyze how effective the application of the Mobile JKN application was in improving the quality of health services during the COVID-19 pandemic.

2 Literature Review

2.1 Electronic Government

E-Government is a concept of a government system which in its governance implementation uses the latest information technology. According to Indrajit, *e-government* is the use of information technology by the government that allows the government to transform relations with the community and in its implementation *e-government* uses the Internet. To carry out government affairs and public service providers so that they are better oriented to community service [11]. The World Bank defines e-government as an effort to utilize and utilize telematics to increase government efficiency and effectiveness, to provide various services to the community better, to provide access to information to the public more broadly, and to be accountable and transparent to the public [12].

Electronic Government is a government system by utilizing *information, communication and technology* (ICT) as a tool to facilitate the process of communication and transactions to citizens, organizations and between government institutions and their staff so as to provide efficiency, effectiveness, transparency and government accountability to the community [13]. In measuring the extent to which E-government can help facilitate in providing competent services to help the community, it can be seen by using E-GovQual indicators, namely, (1) Ease of use, (2) Trust, (3) Reliability, (4) Content and information display and (5) Citizen support (support) [12].

2.2 Electronic Service

The concept of E-Service is a leading application by utilizing the use of information and communication technology (ICT) [14]. According to Rowley, electronic services are defined as actions, efforts or performances whose delivery is mediated by information technology. These electronic services include elements of E-Tailing services, customer support and service [15]. The definition reflects three main components, namely, services, service recipients and service channels (technology). For example, as concerned for public electronic services, public bodies are service providers and citizens and businesses are service recipients. Service channel is the third requirement of electronic service. The Internet is the main channel of e-service delivery while other classic channels are also considered (e.g., telephone, call center, public kiosk, mobile telephone, television) [14].

3 Research Method

In this study, researchers focus on analyzing applications of the JKN Mobile Application in improving service quality during the COVID-19 pandemic. The researcher uses qualitative research methods descriptive. Data was taken from the JKN Mobile Application, Twitter with the *hashtag #MobileJKN,* Online Media, JKN Website and Play Store to support data from researchers who will implement the application of the Mobile JKN application in improving the quality of health services during the COVID-19 pandemic. After that, some of the data collected was processed using the Nvivo 12 Plus Software and then presented descriptively in the observation findings of the Mobile JKN application function, it will be seen its application in improving the quality of health services during the COVID-19 pandemic.

4 Result and Discussion

The electronic government has provided easy access for the community, one of which is in health services in the midst of the widespread spread of the COVID-19 pandemic. In health services where non-COVID-19 patients are also affected, namely, the implementation of social distancing policies, physical distancing and lockdown systems in various regions in Indonesia, making public services difficult to reach. In providing easy services to the public, especially in health services, BPJS Kesehatan has launched an Application and Website-based service provided by the government which plays an important role in preventing the expansion of the spread of the COVID-19 outbreak and aims to make it easier for the community to get health services. The use of the JKN Application and Website has been compiled by BPJS Health so that it can be easily used and reached by all levels of society. The quality of services based on Electronic Government can be measured and seen using indicators, namely, (1) Ease of use, (2) Trust, (3) Reability, (4) Content and display of information and (5) Citizen support (supporters) [12].

4.1 Ease of Use

In the midst of the COVID-19 pandemic, the ease of use of applications provided by the government is no exception The JKN Mobile Application will have a significant impact on the community to assist the community in obtaining the convenience of services needed in the midst of difficult services to reach with the implementation of social distancing policies, physical distancing and the system lockdown, E-government is a solution for the community in getting a service, one of which is in the form of an application. The ease of using the JKN Mobile Application can be seen from how the public responds to the ease of access to the application. The

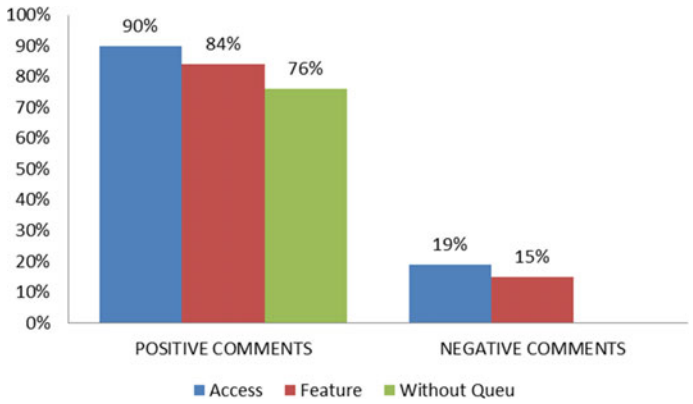


Fig. 1 Community responses in the ease of use of the JKN Mobile Application. *Source* Online Media, Twitter and Play Store

data below will show the public's response to the ease of using the JKN Mobile Application (Fig. 1).

From the data presented above regarding the public's response to the ease of use of the JKN Mobile Application with three indicators that researchers use namely, Access, Features and No Queues, it shows two responses which are different, namely, good response and bad response. The graph above shows that with a percentage of access, 90% received a good response from the community, while 19% received a bad response from the community. Furthermore, regarding the features, where with 84% compared to 15% it can be assessed that the good response from the community is dominant in this indicator. The last thing that the researchers examined was related to good and bad responses in the ease of using the JKN Mobile Application, namely, without queuing with a percentage of 76% getting a good response from the community. It can be interpreted that the percentage of the community's good response is much higher than the community's bad response to the use of the JKN Mobile Application. This shows that the JKN Mobile Application has made it easy for the public to access, the presence of several features and this application is able overcome the accumulation of queues and prevent the expansion of the spread of COVID-19.

4.2 Trust

Use of the JKN Mobile Application which has been structured in such a way that it can be used by all levels of Indonesian society. Before this application can be useful and used by the community, the government must first convince the public to want to take advantage of the application that has been provided. Because trust from

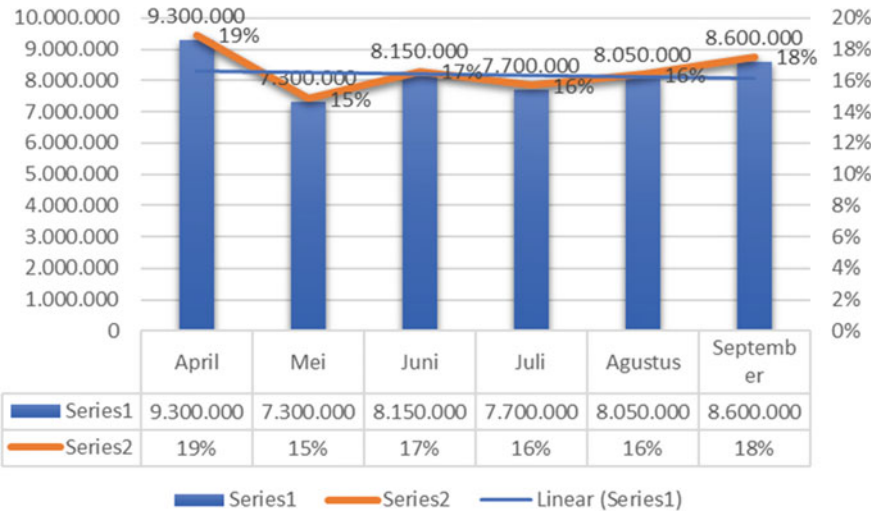


Fig. 2 Mobile application user JKN. *Source* Website JKN

the public will be the first benchmark whether the JKN Mobile Application will be efficient and achieve the targets that have been set in its manufacture (Fig. 2).

Mobile website visitor data from JKN above, in April the number of visitors Website JKN stands at 9.3 million visitors with a percentage of 19%, in May the number of visitors stands at 7,300,000 with a percentage of 15%, in June the number of visitors was 8,150,000 visitors with a percentage of 17%, in July the number of visitors was at 7,700,000 with a percentage of 16%, in August the number of visitors was at 8,050,000 with a percentage of 16%, and in September the number of visitors was 8,600,000 visitors with a percentage of 18%. This shows the level of public trust in using this application is very high. Starting from April to September 2021 the number of visitors to the JKN website is no less than 7,000,000 visitors, the highest visitors are in April with 9,300,000 visitors. These data have shown that the JKN Mobile Application has won the trust of the public. The public already has awareness that the JKN Mobile Application is indispensable during the COVID-19 pandemic in helping to stop and reduce the number of COVID-19 cases in Indonesia.

4.3 Reliability

Mobile JKN Application in placing itself as a facilitator of health service providers, Mobile JKN provides various conveniences in these services, ranging from health administration services, health services at First Level Health Facilities (FKTP) and Hospitals to information and complaints services for JKN-KIS participants. The data below will show the reliability of the JKN Mobile Application (Fig. 3).

Fig. 3 Reliability of the mobile application JKN. *Source* Online Media, Twitter and Play Store

From the data presented above Reliability of the Mobile JKN Application, the results from the Word Cloud show that the dominant words appear, namely, Access and Features. It can be interpreted that many users of the JKN Mobile Application comment on Access and Features. This shows that the JKN Mobile Application provides easy access for the community to obtain health services, besides that the JKN Mobile Application also provides features that can support the community's needs in health services.

4.4 Content and Display of Information

In providing convenience for the public, BPJS also provides online services that utilize chat technology and artificial intelligence, namely, the JKN chat assistant, with this service patients can get a variety of information without having to leave the house (Kompas.com). To prevent the spread of the COVID-19 outbreak while at the same time meeting the community's need for services health, the Mobile JKN Application offers features that reach the community's needs (Fig. 4).

JKN provides 18 features that help people's mobility in daily life and support the community's needs in getting the services they need, namely, information membership, change participant data, bed availability, service registration, premiums, doctor consultations, surgery schedules, relaxation programs for arrears, covered drugs, payment records, participant registration, payments, service history, information and complaints (participants can make complaints in writing or by telephone connected directly to BPJS), JKN info, location, COVID-19 self-screening, screening (Medical History screening contains questions and statements related to Health History).

All of the features provided have received a positive response from the community because they are easy to understand, people just have to choose the health services

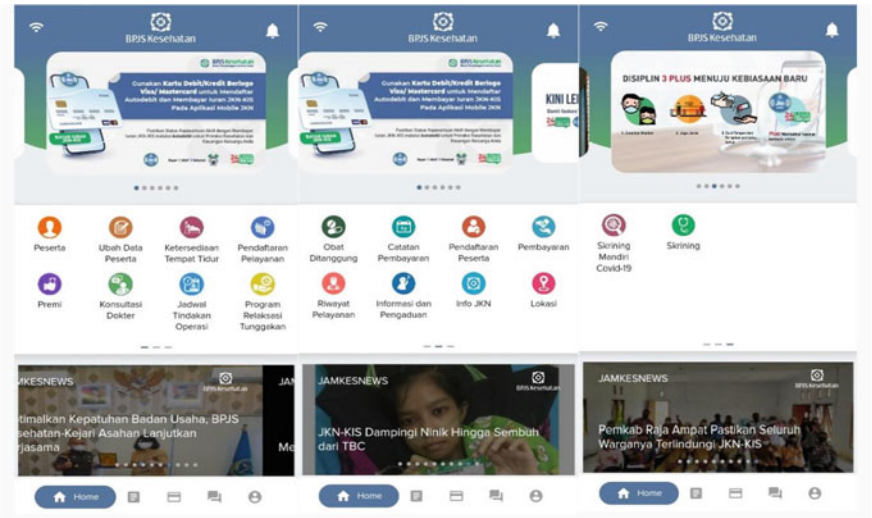


Fig. 4 Contents and information display of the JKN Mobile Application. *Source* Mobile JKN Mobile Application

they need, especially in the midst of the COVID-19 pandemic. As is known, all face-to-face activities, especially in places with large crowds, should be avoided as much as possible during the pandemic. This aims to prevent the transmission of COVID-19 (Kompas.com).

4.5 Citizen Support (Support)

This aspect explains that the community has a guarantee in getting assistance when needed. This aspect relates to the interaction between the community and *customer service* or features that support community complaints if people experience problems or difficulties when using *E-government sites* [12]. Supporting features become a support for the community in accessing an application to provide optimal service to the community. The picture below will show the supporting features to serve the community's problems and difficulties in using the JKN Mobile Application (Fig. 5).

In providing optimal and efficient services to the public through digital-based services, the JKN Mobile Application provides information and complaints features that are useful for serving community problems and difficulties in accessing health and information services. If needed, the community can be helped to convey problems, difficulties and obtain information related to health services through the features that have been provided.

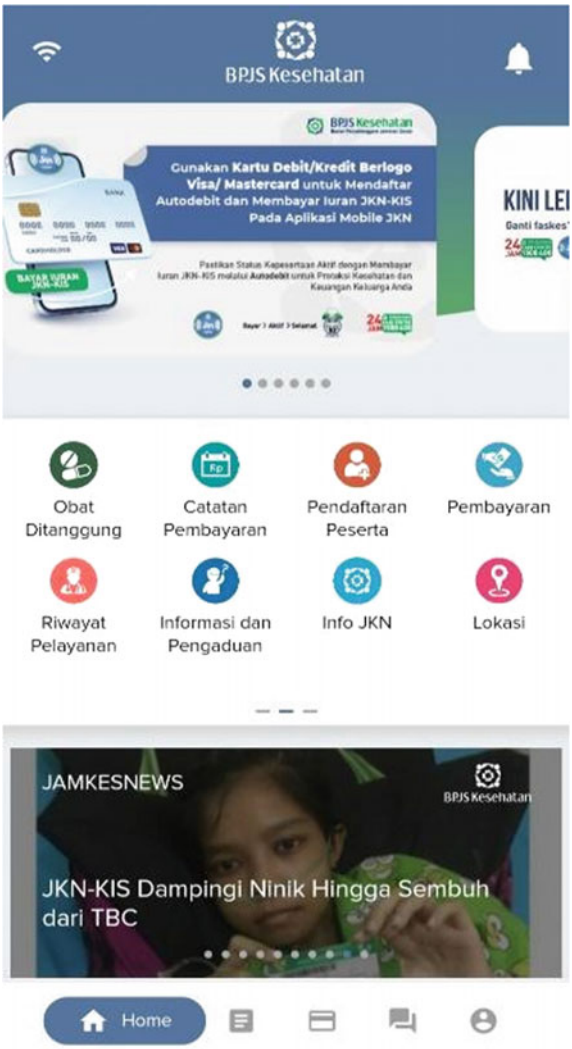


Fig. 5 Supporting features of the JKN Mobile Application. *Source* JKN Mobile Application

5 Conclusion

Based on the analysis using E-GovQual indicators which include (1) Ease of use, (2) Trust, (3) Reliability, (4) Content and information display and (5) Citizen support. The JKN Mobile Application has been able to improve the quality of public services in the health sector during the COVID-19 pandemic. Ease of use, the JKN Mobile Application has made it easier for the public to access with the presence of several features and this application can overcome the buildup of queues and prevent the spread of COVID-19, in terms of trust that the JKN Mobile Application has won the

trust of the public. The public already has awareness that the JKN Mobile Application is indispensable during the COVID-19 pandemic. In terms of reliability, the JKN Mobile Application has provided easy access for the community to obtain health services, besides that the JKN Mobile Application also provides features that can support community needs in health services. Then in terms of Content and Information Display, the features provided in the JKN Mobile Application have received a positive response from the public because it is easy to understand, people just need to choose the health services they need, especially in the midst of the COVID-19 pandemic. Then the last is Citizen Support, the JKN Mobile Application provides information and complaints features that are useful for serving community problems and difficulties in accessing health services and required information. In the future, the JKN Mobile Application is expected to be better and more reliable to assist the community in obtaining health services.

References

1. Agung IM (2020) Memahami Pandemi Covid-19 Dalam Perspektif Psikologi Sosial. PsikobuletinBuletin Ilm. Psikol. 1(2):68–84. [Online]. Available http://ejournal.uin-suska.ac.id/index.php/Psikobuletin/article/view/9616/5058
2. Hasti I, Nurmandi A, Muallidin I, Kurniawan F, Salahudin (2022) Pros and cons of vaccine program in Indonesia (Social Media Analysis on Twitter), vol 319(84). Springer International Publishing
3. Febty IM, Nurmandi A, Muallidin I, Kurniawan D, Salahudin (2022) The effectiveness of social resilience in Indonesia, vol 319. Springer International Publishing
4. Effendi N, Widiastuti H (2014) Jurnal Kesehatan. J. Kesehat. 7(2):353–360. https://doi.org/10.24252/kesehatan.v7i2.54
5. Felicia FV, Suarca IK (2020) Pelayanan Imunisasi Dasar pada Bayi di Bawah Usia 12 Bulan dan Faktor yang Memengaruhi di RSUD Wangaya Kota Denpasar Selama Masa Pandemi COVID-19. Sari Pediatr. 22(3):139. https://doi.org/10.14238/sp22.3.2020.139-45
6. Taufik T, Warsono H (2020) Birokrasi baru untuk new normal: tinjauan model perubahan birokrasi dalam pelayanan publik di era Covid-19. Dialogue J. Ilmu Adm. … 2(1):1–18. [Online]. Available https://ejournal2.undip.ac.id/index.php/dialogue/article/view/8182
7. Haryadi T, Nurmandi A, Muallidin I, Kurniawan D Implementing "SIREKAP" application based on election for improving the integrity of election administrators and increasing, vol 1. Springer International Publishing
8. Wulanadary A, Sudarman S, Ikhsan I (2019) Inovasi Bpjs Kesehatan Dalam Pemberian Layanan Kepada Masyarakat : Aplikasi Mobile Jkn. J Public Policy 5(2):98. https://doi.org/10.35308/jpp.v5i2.1119
9. Widyasari SANP, Yasintha PN, Wirantari IDAP (2019) Implementasi Program Mobile Jkn Dalam Memberikan Pelayanan Pengguna Kis Di Kantor Bpjs Kesehatan Kabupaten Karangasem. Citizen 1(1):1–9
10. Wahyudi W, Nupus H, Pascasarjana P, Manajemen M, Sultan U, Tirtayasa A (2021) Meningkatkan Intention To Use Aplikasi Mobile JKN Melalui Perceived Usefulness Dan Attitude Towards Use Sebagai Variabel Intervening Pendahuluan. 5(1):98–116
11. Widodo N (2016) Pengembangan e-Government di Pemerintahan Daerah Dalam Rangka Mewujudkan Smart City (Studi di Pemerintah Daerah Kota Malang). J. Ilm. Adm. Publik 2(4):227–235
12. Maulani W (2020) Penerapan Electronic Government Dalam Peningkatan Kualitas Pelayanan Publik (Studi Kasus Program E-Health Di Kota Surabaya). As Siyasah 5(2):44–54

13. Aminudin N, Putra DAD (2014) Langkah-Langkah Taktis Pengembangan E-Government Untuk Pemerintahan daerah Kabupaten Pringsewu, vol 3
14. Buchari RA (2016) Implementasi E-Service Pada Organisasi Publik Di Bidang Pelayanan Publik Di Kelurahan Cibangkong Kecamatan Batununggal Kota Bandung. Sosiohumaniora 18(3):225. https://doi.org/10.24198/sosiohumaniora.v18i3.8762
15. Naqibah LS, Cikusin Y, Abidin AZ (2021) Implementasi Kebijakan Pelayanan Administrasi Kependudukan Berbasis E-Service Dinas Kependudukan dan Pencatatan Sipil Kabupaten Rembang. J. Respon Publik 15(9):22–30

The Dynamics of Cyber-Activists in the Digital Era of Papua, Indonesia

Yuspani Asemki, Achmad Nurmandi, and Isnaini Muallidin

Abstract The Papuan conflict is classified as an internal conflict, which occurs within sovereign state. This conflict cannot be separated from the community's dissatisfaction with the Indonesian government's services to the Papuan people. The Papuan conflict has its roots in a long historical event, starting with the history of the entry Dutch colonialism, and the Dutch ending in 1969, in the same year Indonesia entered through the Act of Free Choice. But, some of the Papuan people and some international elements think it is not finished yet. The purpose of this research is to know the dynamics of cyberactivism in handling cases of conflict in Papua, Indonesia. The method used is a qualitative method using a descriptive approach that produces data on the dynamics of cyber-activists in the digital era of the conflict in Papua, Indonesia which will produce results data. Data obtained were om documents, journals, theses, and news. Data are managed through Nvivo. The data processing mechanism describes, analyzes phenomena, social activity events, attitudes, beliefs, perceptions, and human thoughts that are conveyed through the media and other documents that have been digitized. In this study, the author uses the theory of cyberactivism, as an analytical material to answer the title that will be discussed regarding the dynamics of cyber-activists in the digital era of Papua, Indonesia. This theory is used as a basic reference material to find out the use of social media in Indonesia, more specifically the problem of the dynamics of the conflict in Papua, Indonesia.

Keywords Dynamics · Cyberactivism · Papua · Indonesia

1 Introduction

The dynamics of the Papuan conflict have been on the international radar screen since Indonesia became an independent country in 1945. Since the handover of Papua to Indonesia in 1962, low-intensity military conflicts have developed. Most

Y. Asemki (✉) · A. Nurmandi · I. Muallidin
Department of Government Administration, Jusuf Kalla School Government, University of Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
e-mail: yuspani.asemki.pcs21@mail.com

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_35

Papuans believe that their right to self-determination is sacrificed at the altar of geopolitics. Then, formed the Free Papua Organization (OPM) on May 6, 1973 to fight the Indonesian government. In this important moment, the Papuan people always remember and want their independence. While the Central Government has taken various approaches, such as delegating authority to the Papuan provincial government, this has not had a positive impact on conflict resolution for the Papuan people, instead increasing the conflict. Papuans are treated unfairly by the government, thus causing geopolitical dynamics that occur between Indonesia and Papua. This dynamic, of course, the media is involved in reporting the problems that occur between the two sides of Indonesia, Papua [1].

The Indonesian government (Mahfud MD) said that the political dynamics of the Papuan conflict have always been forgiven by the KKB (Armed Criminal Group), quoted from the YouTube channel of the Coordinating Ministry for Political, Legal, and Security Affairs, Thursday (29/4/2021). This is contrary to the actual behavior of the state toward the Papuan people. The government also sent the Indonesian National Armed Forces (TNI) massively for security reasons and to maintain the sovereignty of the Indonesian state; however, this traumatized the Papuan people with the presence of the TNI. One of the main reasons for being secure is to eradicate the KKB and to secure the Papuan people, especially around conflict areas, but not handling them properly. In fact, the presence of the TNI in Papua has traumatized the Papuan people by remembering the military operations of the 90s. Mahfud MD also said that the KKB also caused the destruction of vital strategic objects to the environment, public facilities, international facilities with ideological, political, and security motives [2].

The dynamics of the Papuan conflict, of course, the media take part in publicizing all the problems that exist in Papua without taking sides with any party, be it the Indonesian government or the Papuan people. Contribution The dynamics of cyber-activism regarding Indonesia's conflict with the Free Papua Organization (OPM) are currently more in favor of the Indonesian government and cover up all the problems that have been rotting for a long time in the Papua region from 19,962 to the present. We always get different views from the dynamics of cyber-activists who publish news that is always broadcast through TV stations, newspapers, and other online media to respond to the real problem of Papua. We follow the dynamics of cyber-activists that are currently being played by siding with one party, namely the Indonesian government and never raising issues that are actually happening in Papua. There are several media that try to raise the issue of the dynamics of the Papuan conflict; however, this is always covered up and threatened by the Indonesian military and diverts issues that are not right on target to the realities and conditions that actually occur in the Papua region, publicizing events in Papua such as murders, rape, theft, and others are also not in accordance with clear sources of information and even more lead to opinions that provoke the Papuan people [3].

Torture and state-sponsored peacebuilding encapsulate the essence of many of Indonesia's current conflicts. Papua in particular provides a thought-provoking example of the complexities and complexities of building peace in the midst of protracted conflict and violence. Indonesia is a unique case because the government has done nothing in dealing with the OPM strategy. The circulation of videos of

torture carried out by members of the TNI in Papua can actually be considered a form of netwar pattern that benefits the OPM, but it turns out that the Indonesian government only gave a statement that it would investigate the truth of Indonesia's response in the form of neglecting the OPM network strategy based on profit-and-loss considerations, in the sense that the pattern neglect is more profitable for the Indonesian government than the pattern of resistance or signaling back to explain how the constructed news text represents counter-propaganda. The use of media is currently widely used by users to propagate issues of the Papuan conflict by selecting issues or highlighting aspects of a certain reality that exists in Papua [4].

Some of the results of the discussion above show that the dynamics of using cyber-media to deal with the conflict between Papua and Indonesia are not neutral toward publicizing events that occur in Papua. There are so many media that lead to opinions that make the people of Papua not accept because they publish events that are not in accordance with what is experienced by society today, and publish news from one side only and blame one side, namely the Indonesian government often blames the OPM and develops news. Provocateur to the duas community. The Indonesian government, OPM, and activists in Indonesia currently have no political solutions offered to the public to be resolved with dignity. This matter, Pastor Neles Tebay once offered to Indonesia for dialogue with the authorities and was mediated by a third party; however, this was considered negative by the Indonesian government [5].

The dynamics of today's saber activists to see the prolonged conflict in Papua, of course, we will look at several aspects, namely why Papua is always in trouble with the problem of Papua's integration into Indonesia, regarding social, economic, health inequality, and several other factors that we can both share finish well. But, Indonesia is currently not looking at the things I have mentioned above. Indonesia is currently focusing on maintaining the security and sovereignty of the Republic of Indonesia and sending a very massive military to the Papua area. In addition to sending the military to Papua, there are several generals who take advantage of the situation and conditions to invest in illegal mining with military force in Papua, such as the Vice Regent's Bloc in Intan Jaya district, Palm Oil in South Papua, and several other illegal mines in the Yahukimo and Bintang Mountains areas encouraged by the Indonesian military to make it easier to operate in several illegal mining companies in Papua [6].

2 Literature Review

2.1 Cyberactivism

Cyberterror began to be known in 1995 when researchers used (electronic Pearl Harbor) as the first warning for the United States. Since then, researchers to policymakers as well as military officials have highlighted how terror can evolve along

with the development of technology. The climax occurred when the events of 9/11 hit the World Trade Center in 2001, and a group of hackers called G-Force Pakistan acknowledged their contribution to the tragedy. There are so many academics who try to formulate the definition of cyber-theory to study in order to find a solution and be able to resolve the conflict problem in Papua, Indonesia; however, no one has really created conducive cyberspace for activists regarding the conflict in Papua, Indonesia. Pollitt in his article entitled cyberterrorism is the theory of cyber as a planned activity with a political motive to attack information and technology systems that aim to attack a group of targets (Pollitt 1998) [7].

The cyber-theory is applied to be accessed more freely through the actual use of media regarding situations and conditions surrounding media use. The cyber-theory also aims to facilitate the right of opinion for people who are under repressive government regimes. Ironically, now the Darknet seems to be a rat's way that is loved by conflict. So, the description of the various studies above confirms once again how cyber-terror continues to evolve, and its impact is expanding to the realm of Internet freedom. This expansion does not make terror in cyberspace more flexible but brings terrorists closer in everyday life without us realizing it [7]. Cyber-theory invites us to be able to use social media as a medium of communication and to share information, both good and bad, for our common consumption together. This theory also makes the basis for detecting the use of media among the public, both the political elite and the bureaucratic level. The author uses this theory as a basis for knowing the dynamics of the use of social media regarding the conflict in Papua, Indonesia [7].

3 Methodology

This article uses qualitative research methods and uses a descriptive approach to analyze and generate data on the dynamics of conflict activism in Papua, Indonesia. According to Sugiyono, qualitative research is research where researchers are placed as key instruments; data collection techniques are carried out in combination, and data analysis is inductive to produce the intended results [8]. This study aims to examine and analyze the dynamics of digital activism in the Papua, Indonesia conflict and to find out the opinions and views of the Papuan people regarding the Papuan conflict on social media as well as issues that are developing among the wider community regarding Free Papua and to map out organizations that voice Papua Papua based on with the problem of humanitarian crisis and social and cultural gaps. Data are collected through Scopus and several other media, such as mass media, journals, theses, and theses and other documents that have been digitized to answer the dynamics of cyber-activists in the digital era regarding the issues of conflict in Papua, Indonesia. The data are collected through Scopus from various Web sites and news and is managed through the Nvivo application which produces primary and secondary data to find out the dynamics of cyberactivism regarding the conflict in Papua, Indonesia. The theory used by the author for this article is to use Activism Theory to be able to pay for the dynamics of saber activism which is negotiated

by local elites and several other groups regarding the issue of the conflict in Papua, Indonesia.

4 Results and Discussion

Dynamics of the development pattern of activism in the digital era in the use of social media in Indonesia, more specifically on the issue of the conflict in Papua, Indonesia. The role of the mass media that is not in favor of any institution or agency, in publishing news, event, or information conveyed through the mass media to the public to be consumed by the public, besides that the media must stand-alone or independent. News sources published by the media must use clear and accurate news sources so that they can be trusted by the public. Because the news becomes the benchmark as the main basis for the occurrence of an event at the scene. Processing an event properly and accurately so that it becomes quality, neutral, and reliable news for the community, especially news about the Papuan conflict which is always brought up by several media as well as a source of reference for the people of Papua and outside Papua. Besides that the management of irresponsible public opinion is also increasingly prevalent among the wider community regarding the issue of the Papua conflict, Indonesia [9].

While the Indonesian government has also responded to the issue of the Papuan conflict on social media, there has always been a controversy between pro-independence organizations. Likewise, pro-Papuan independence organizations on social media, so that the dynamics of the political problem of the Papua–Indonesian conflict have long been happening on both sides. The pattern of the dynamics of Indonesian and Papuan politics has been going on for a long time among the Papuan people. However, there is no political offer offered to the public as a democratic solution to resolve the conflict in Papua, Indonesia as a way of peace. There are several solutions the Indonesian government offer to the people of Papua, as a way to resolve the conflict in Papua, Indonesia, such as Law Number 21 of 2021 concerning Special Autonomy for the Papua Province, educational opportunities for Papuan youth, job opportunities at the provincial and ministry levels. However, this does not have a positive impact on the Papuan people. Humanitarian Activist, Father Neles Tebay through the Papua Peace Network (JDP) offered to Government for a peaceful dialogue, and the dialogue was brokered by a third party, namely the United Nations. However, this was rejected by the Government of Indonesia with a negative perception toward JDP and the Papuan people of the Papua-Indonesian conflict, as follows.

4.1 The Dynamics of Papuan Conflict Cyberactivism

Recently, the Papuan conflict has become a hot topic of discussion at the local, national, and even international levels; therefore, the author looks at the dynamics of cyberactivism which acts as a mediator to be able to inform the Papuan conflict issues to the wider community, especially within the state of Indonesia so that it can know the problem in Papua. Because the Papua conflict that has been going on for a long time since the 1960s until now cannot be finished. Furthermore, in this discussion, we focus on the dynamics of activism to be able to map the issues of the conflict in Papua, Indonesia, in terms of handling activism, an issue that is often discussed by activists in handling the conflict in Papua, Indonesia. Based on the results of the study, the authors present the results of the dynamics of cyberactivism in the conflict in Papua, Indonesia as follows (Fig. 1):

Based on the data obtained by the author through the Nvivo application regarding indicators the dynamics of cyberactivism regarding the conflict in Papua, Indonesia. The narratives that appear most frequently in discussions of cyberactivism are about activism issues, meaning that the Papua and Indonesia conflict cases are widely discussed on social media compared to the implementation of real handling in the field in accordance with what is desired by the wider community. This is followed by indicators of activism in the Papua conflict, Indonesia, where activism accompanies the issue of the Papuan conflict as an object that turns a cyberactivism business in the conflict in Papua, Indonesia.

This is followed by a cyberactivism indicator where this indicator focuses on the dynamics of conflict activism in Papua, Indonesia. In the dynamics of saber activism, there are many discussions that are discussed by actors who are members of the government and non-government which are discussed with their respective interests. This actor is not only one party but, involved in it many parties are involved such as pro-independence groups, political elite groups, community leaders, traditional

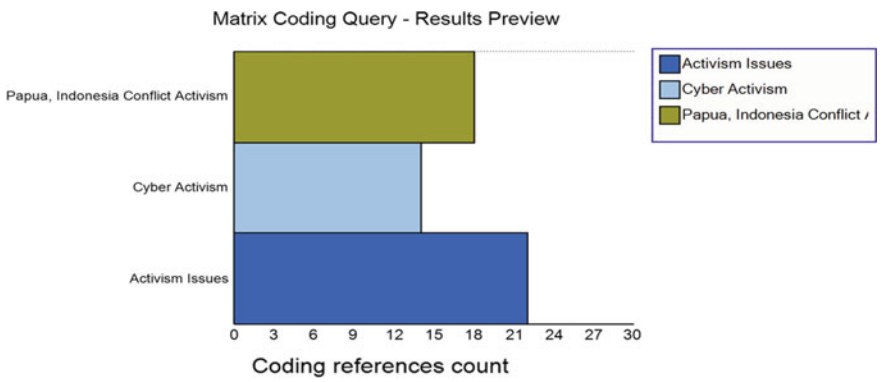


Fig. 1 Matrix coding Nvivo 12 plus, indicators of dynamics of cyber-activism in the conflict in Papua, Indonesia. *Source* Compiled by the Author

leaders, and Papuan activists who voice the issue of human rights violations as concerned with their personal interests, and there are also those who are concerned with the public interest such as raising the issue of the humanitarian crisis, social inequality, economy, culture, and several other issues.

4.2 Actors of the Dynamics of Papuan Conflict Activism, Indonesia

Actors of the dynamics of Papuan conflict activism, in Indonesia, are people, groups, organizations, communities, and even the greatest countries in the conflict in Papua, Indonesia. Regarding the Papua conflict, Indonesia is the basis of the problem that must be handled seriously by the Government of Indonesia and Papua so that it can be resolved peacefully and with dignity. Looking at some of the actor's responses, both actors as political elites, chords as government, and individual organizations are involved in the conflict in Papua, Indonesia. Furthermore, the Papua conflict, in Indonesia, shows that the problem is not only one issue that must be responded to by elements of the Papuan community but several issues at issue such as the ideology of Papua Merdeka, social inequality, a cultural economy that must be considered by leaders, activists, academics, and the wider community in Indonesia in general and in Papua in particular. Here, the authors present the results of research that have been obtained from several studies through Nvivo, more specifically conflict actors in Papua, Indonesia, as follows (Fig. 2):

Based on data obtained through Nvivo regarding the indicators of Actor Dynamics of Conflict Activism in Papua, Indonesia. The first indicator shows the Central Government, which means, a policy regarding the handling of the Papuan conflict problem, Indonesia, is centered on the Central Government; in handling this policy, it is always in conflict with the wishes of the community so that the Papuan people

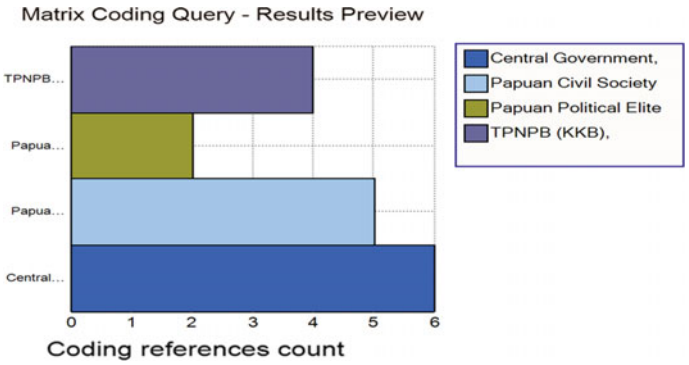


Fig. 2 Matrix coding Nvivo 12 plus, indicators of actor dynamics of conflict activities in Papua, Indonesia. *Source* Compiled by Author

reject the implementation of the policy on dealing with the Papuan conflict issue. On the other hand, the Central Government imposed the policy so that hardliners who did not accept the implementation of the policy turned toward creating conflicts such as the horizontal conflict, the Indonesian conflict with the pro-independence West Papua. This is followed by other indicators that discuss what are the most dominant actors in it that can be discussed regarding the Papua conflict, Indonesia is as follows: First, the actor that is most often discussed regarding the Papua issue, Indonesia is the Central Government; within the Central government, there are several institutions or related elements that the author acknowledges in the structure of the Indonesian government such as ministry institutions (law and human rights), Papua regional government, generals, former presidents, and several independent institutions that work with the Central Government to discuss issues of conflict in Papua, Indonesia.

Furthermore, the indicators of civil society, in fact, the civil society living in the Papua region wants peace in the Papua region so that the people can be calm and peaceful in carrying out all activities in the environment where they live. The Papuan people want peace, but there are several related actors who have personal interests that cause conflict in their environment, as a result of the conflict, people flee everywhere to the forest. The Papuan people hope that this conflict will soon be resolved properly with dignity and besides that they hope that violations of human rights (HAM) can also be resolved properly, honestly, fairly, and hope that the actors involved in it can be brought to justice by the state because we are here in the rule of law [10].

Furthermore, the Toko Papua and KKB (TPN/PB) indicators are more or less the same in the discussion, meaning that both parties discussed several political issues such as the ideology of an independent Papua, human rights violations, social inequality, exploitation, culture, and Papua's depleted natural resources illegally in addition to the state's illegality; it is driven by the massive militarism in Papua. Toko Papua is classified into two groups, namely the government group and individual groups/institutions that are pro-independence Papua. Groups or shops that talk about the issue of an independent Papua such as the West Papua National Committee Organization, the West Papua Liberation Army Group, Wi-Fi AI, and several individuals who talk about an independent Papua such as Beny Wenda, Pastor Beny Giay, Filep Karma, and several Papuan figures who discuss the issue of Free Papua and the humanitarian crisis.

Meanwhile, shops from the government side have voiced several issues that are often discussed such as issues of human rights violations, welfare issues, social inequality, division of authority at the provincial and district level, exploitation of Papua's natural resources without the knowledge of the owner's rights, and several issues or problems discussed by local Papuan shops or elites. Papuan shops/political elites who have been included in government institutions are often referred to in response to the Papuan conflict issues, especially regarding the issue of Papuan independence which has triggered a prolonged conflict that has not been resolved until now. Besides that the Papuan regional government is also not firm in handling the Papua–Indonesian conflict. The Papuan local government must always participate in the policies of the Central Government (Jakarta) to implement its governance. While

the Papuan political elite always follows the will of the Central Government and does not pay attention to the basic rights of the Papuan people, this causes horizontal conflicts within the political elite itself [11].

4.3 Issues of the Dynamics of Conflict Activism in Papua, Indonesia the Issue

The problem of the dynamics of conflict activism in Papua, Indonesia is how we look at the dynamics of politics built by related actors, which has been explained in the previous indicator that how actors play or develop issues that are built structured in accordance with the respective issue areas that are built. Looking at the previous indicators, we can map out that the political dynamics of the Papua–Indonesian conflict have always been questioned by several developing issues such as the issue of human rights violations, racial discrimination, exploitation of natural resources, social inequality, pro-independence Papua, and other political elite interests played by actors related to actors that have been described above. Based on this brief description, the author presents the results of research carried out by the Nvivo application to be able to find out the results regarding the dynamics of the activism issue of the Papuan conflict, Indonesia. In the dynamics of cyberactivism, the issues that are questioned by their pre-political elite, political elites, and even several other figures who always raise the issue of the conflict in Papua, Indonesia. Here, the authors present the results of research that have been obtained from several studies through the Nvivo application, more specifically on the issues discussed in the cyberactivism of the Papua, Indonesia conflict, as follows (Fig. 3):

Indonesia conflict conditions that have not been resolved by the Central Government for a long time; besides that there are several problems, namely human

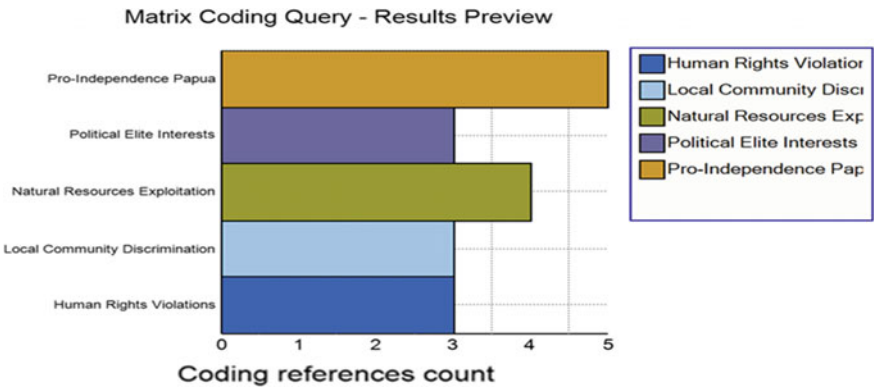


Fig. 3 Matrix coding Nvivo 12 plus, indicators of the dynamics of conflict activities Papua. *Source* Compiled by Author

rights violations, illegal exploitation of natural resources, social, economic, cultural inequalities, and several other problems, the Papuan people choose to support their Papuan identity. However, there are some political elite groups who prioritize their personal interests to a campaign that the Papua region has officially become part of Indonesia and the Unitary State of the Republic of Indonesia Price Mati but what the Papuan people really want is the independence of West Papua. The next indicator is the exploitation of Papua's natural resources, with mining companies in Timika, oil in Sorong and Kaimana, coal in Maybrat, gold in the Jayawijaya Mountains, and Intan Jaya, Block Wabu, etc. In addition, there are oil palm companies in Merauke, Arso, Sorong, and several timber companies operating in the Papua region. Several companies operating in the Papua region sometimes operate illegally without a business license (SIU) to take land from companies that are operating or are currently under operation. To take over the land or territory, the company always contradicts or clashes with the local community because it enters without the knowledge of the community or the local tribal chief. Sometimes, to give reasons to the community so that the community believes, the company owners claim that there is already a letter of approval from the local tribal chief to operate the land as a mining site, in fact, there is no letter of approval from the local tribal chief, even though there is one, but it is only on behalf of the local tribal chief for the sake of the company's net [12].

Next is the indicator of human rights violations. This issue cannot be separated from, with the continued exploitation of Papua's natural resources, both legal and illegal companies operating in Papua. In addition, the company also does not want to prosper the residents in the surrounding area and subsequently takes the company's land forcibly and illegally. Thus, massive conflicts and gross human rights violations have arisen to this day. Meanwhile, the state also intervened in the conflict because the actors were regional leaders of the state, so the State Government ordered to send the military to Papua with the aim of protecting the territory of the Republic of Indonesia, but on the other hand, they also guarded natural resources (SDA) which were being exploited officially or illegal in the Papua region; for that reason, it also caused many casualties and increased serious human rights violations in Papua [13].

Next up is the pro-independence indicator. This indicator is also inseparable from the problem of continued exploitation of natural resources; however, it does not pay attention to the Papuan people in the surrounding environment, human rights violations, and some basic things that make an impetus to advocate for an independent Papua. While the grassroots communities in the Papua region all say they want to be independent themselves, however, the Indonesian government is adamant not to let go because Papua contains abundant natural resources. Besides that to control the natural products, some political elites also take advantage of their interests to take advantage of this moment for their personal interests by building issues that do not agree with or oppose the real wishes of the community.

5 Conclusion

Based on the results of the discussion above, the writer concludes that all Papuan people want that the Central Government and the Papuan Government to immediately resolve the dynamics of the Papua–Indonesian conflict, which has been long and unfinished, as well as causing serious human rights violations in Papua. In resolving the Papua conflict, Indonesia will use peaceful means such as peaceful dialogue with pro-independence organizations and the Indonesian government mediated by a third party, namely the United Nations. With gross human rights violations, social, cultural, economic disparities, the high poverty rate in Papua makes the Papuan people do not believe in the existence of the State of Indonesia in the Papua region, the existence of the State of Indonesia in Papua has been 53 years, but there has been no significant change for the Papuan people, even the people of Papua. Papua considers Indonesia's presence in Papua considers it to be increasingly backward. This makes the Papuan people more confident in wanting the independence of the West Papuan nation and the spirit and confidence that Papua will be free from the hands of the Unitary State of the Republic of Indonesia.

References

1. Pichai N (2014) Peace solutions in Papua land (burying human rights tragedies and finding ways of peace) 1 Natalius Pigai 2, peaceful solutions in Papua (burying human rights tragedy and finding a way of peace). 11:23–38 [Online]. Available: https://media.neliti.com/media/publications/72813-ID-solution-peace-di-tanah-papua-bury-tra.pdf
2. Yahya AN (2021) The government's reason for establishing KKB in Papua as a terrorist organization. Kompas.com [Online]. Available: https://nasional.kompas.com/read/2021/04/29/13263971/alasan-government-tetapan-kkb-di-papua-organization-terrorist?page=all
3. What was the reason for Gus Dur not to prohibit the raising of the morning star flag on page all—Kompas
4. Noor G et al (2019) The Indonesian government's strategy against MSG countries …|Putra, Legion Suko, Madjid|31, vol 5, no 2. University of Defense, pp 31–44 [Online]. Available: http://jurnalprodi.idu.ac.id/index.php/PA/article/view/409
5. Tebay N (2016) The transformation of the Papuan conflict. J Relig Culture 12(April):82–106
6. Sugandi Y (2008) Conflict analysis and policy recommendations regarding Papua. Friedrich Ebert Stiftung, pp 1–30
7. Wijayanto W, Sardini NH, Elsitra GN (2021) Creating a conducive cyberspace for anti-corruption activists. Integrity 7(1):179–196. https://doi.org/10.32697/integrity.v7i1.732
8. Furqon MA (2019) The dynamics of resilience in widows (a case study on a widow whose spouse died in middle age in the hamlet of Plumpung Rejo, Karang Tengah Village, Kediri). J Online Int Nas 7(1), January–June 2019 Univ. 17 August 1945 Jakarta, 53(9):1689–1699 [Online]. ISSN 2502-3632 ISSN 2356-0304. Available: www.journal.uta45jakarta.ac.id
9. Filep Karma no more special autonomy for Papua, we want independence—I Papua 2021-10-16 22_03_08Z
10. Brata AG (2008) Regional expansion in Papua: community welfare VS elite interests. Simp Nas Ris Wisdom Icon 1–9
11. Yanampa Y, Sulistyaningsih T, Norman A (2017) Cooperation relationships with local elites abstraction. Semin Nas Prod Degree 1: 1069–1080

12. Taum Y (2019) Violence and conflict in Papua: root causes and strategies for overcoming them. J Res 19(1):1–13
13. Filep Karma says OAP wants welfare—I Papua 2021-10-16 21_59_23Z

OMNIBEE: Autonomous Omnidirectional Robot for Service Robotics Applications

Víctor H. Andaluz, Christian P. Carvajal, Jenny Granizo, José Varela-Aldas, Luis E. Proaño, and Danny Pérez

Abstract This paper proposes the construction of the software and hardware of an omnidirectional mobile robot with the ability to receive linear and angular velocity commands to move from one place to another. In the prototype can be used controllers suitable for their configuration, capable of performing movements of the omnidirectional platform for the execution of specific tasks, the robot will also transmit the current speeds to have a feedback of any control, i.e., a control law for the execution of positioning, trajectory tracking, road tracking, among others.

Keywords Omnidirectional robot · Controllers · Tasks · Feedback · OMNIBEE

1 Introduction

The area of robotics has evolved, with the purpose of satisfying the needs of users. Nowadays, robots can be implemented in the industrial area and in environments that interact with humans [1, 2]. In these areas robotics has demonstrated its high efficiency, velocity, precision, is not tired and coordinates perfectly with the rest of robots and machines involved in a process [3]. Its current features allow you to easily position yourself in the desired place in a sub-hierarchical way, locate parts

V. H. Andaluz (✉) · L. E. Proaño · D. Pérez
Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador
e-mail: vhandaluz1@espe.edu.ec; victorandaluz@uti.edu.ec

C. P. Carvajal
Instituto de Automática, Universidad Nacional de San Juan, San Juan, Argentina
e-mail: cpcarvajal@inaut.unsj.edu.ar

J. Granizo
Universidad Nacional de Chimborazo, Riobamba, Ecuador
e-mail: jgranizo@unach.edu.ec

V. H. Andaluz · J. Varela-Aldas
SISAu Research Group, Universidad Tecnológica Indoamérica, Ambato, Ecuador
e-mail: josevarela@uti.edu.ec

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_36

with which you have to work, communicate with other equipment, be programmed using effective computer tools, etc. [4, 5].

The applications of mobile robots exceed human ability to solve and meet objectives that humans can be complicated by personal security or other factors. The benefits, advantages and applications when using robots have been transformed into user needs since using robots complicated tasks are fulfilled without human intervention, even without robotics thousands of people were not able to perform activities by, e.g., Walking, flying, manipulating objects, transport tasks, among others [6, 7]. In the area of mobility robots used to perform activities in unstructured environments are of various types using legs, wheels, aquatic robots, aerial, some of them even have built-in manipulators in their structure.

The implementation of robots capable of moving in a non-structured environment [8], within the development of robotics can be controlled autonomously or tele operated, either for industrial applications such as assembly of components, production of devices, health areas, Manipulation of objects, for service robotics, among others [9–11]. In the case of [12] using unicycle type robots to move an object from an initial position to a desired position by applying position, path and path planning controllers, the disadvantage of this type of mobile robot is that it only has a linear velocity vx and one angular velocity ω for translation. The omnidirectional robot presented in this work considers mecanum type wheels, which allow frontal linear displacement, lateral displacement and angular displacement [13–15].

2 Omnidirectional Robot

Taking into account the configuration of the omnidirectional robot, different components such as motors, mechanized wheels, encoders and different parts are used which will help the platform to move in every direction in the plane, the generated force of the omnidirectional robot to produce the movement is given by the four 4 motors that are coupled directly to each omnidirectional wheel of mecanum type. In the design of the structure takes into account some properties that the robot must have, i.e., the weight it supports, the distribution of mechanical and electronic elements, and geometry suitable for its displacement [16]. In the design of the structure is considered the strength and stability, which is why the platform has a rectangular shape of a prism, using aluminum.

Through a simulation, the behavior of the structure is determined, which is subjected to a real force that is applied in the places where it has the highest possible load. Figure 1 a Finite Element Analysis (FEA) [17] analysis is performed, which calculates the Tensions or deformations and pressures to which it is subjected, Fig. 2 indicates the displacement of the material when applying a force (blue color), i.e., that the deformations caused by the forces applied are very small, so the rigidity caused by the load is not taken into account.

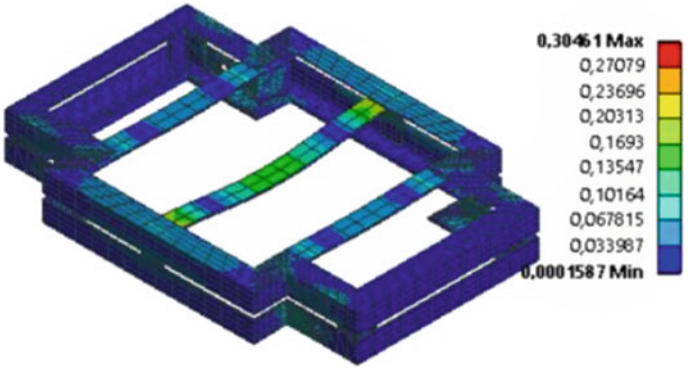


Fig. 1 Analysis of tension of the structure

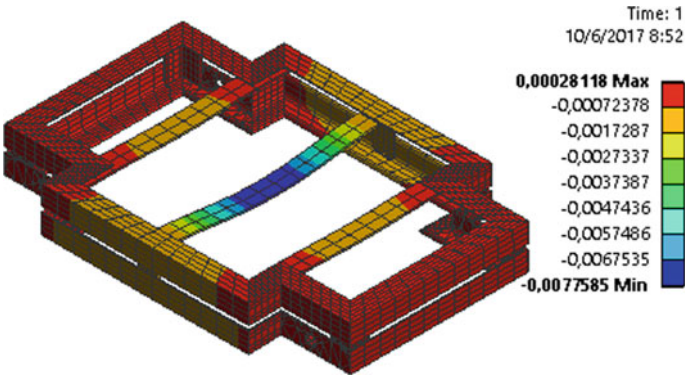


Fig. 2 Displacement of tension of the structure

The culminating design of the omnidirectional robot structure consists of a few printed parts to shape the platform without interfering with the size and set dimensions since the media is PLA with negligible weight.

3 Hardware and Interface Implementation

The hardware of the omnidirectional robot is comprised four stages, i.e., consists of sub electronic structures that allow the robot to move within the plane X, Y and rotate around the axis Z. The stages taken into account for the construction of the prototype are: (i) power system (ii) control system, (iii) actuators and sensors stage, finally (iv) the wireless communication stage, see Fig. 3.

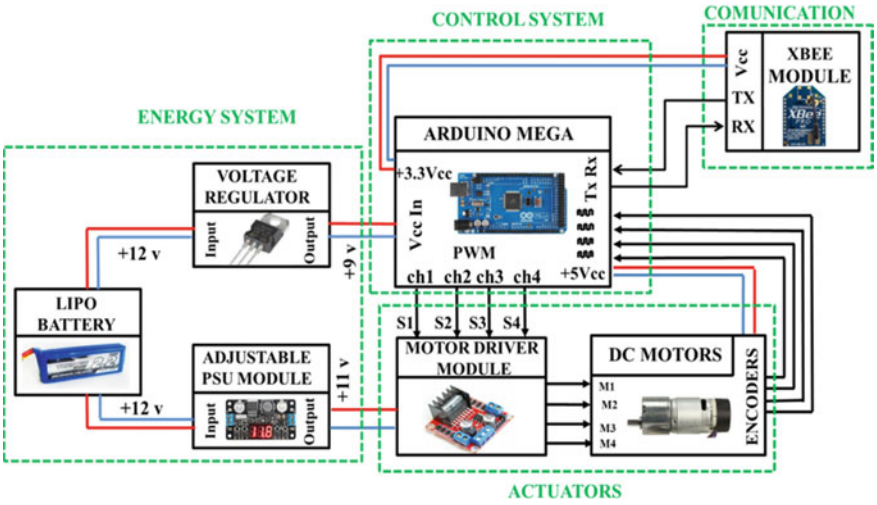


Fig. 3 Mobile platform hardware

Power system: The omnidirectional robot power system consists of a 4-cell LIPO (Lithium-Polymer) battery with a supply current of 2.2A, enough to provide the necessary current to power the entire robot system; the supply voltage for the motor drivers is 11 V, i.e., it has a voltage regulator that adjusts to the maximum acceptable value for the motors. For the control board voltage control, a 9 V regulator source was used; the control board supplies 5Vdc and 3.3Vdc for the sensors power supply and for the communication module, respectively.

Control system: The robot control system is composed of an Arduino Mega board which performs various tasks such as, *e.g.,* receive the velocities corresponding to each wheel of the platform obtained from the calculation of the velocities lines and angular velocity to be applied to the robot to move according to the application desired by the user. This controller compensates the dynamics of the omnidirectional robot, where $W_d(t)$ represents the desired angular velocities for each robot wheel and $W_s(t)$ the measured angular velocities of the robot by the sensors, i.e., the speed errors are defined by: $\tilde{W}(t) = W_d(t) - W_s(t)$ this error is compensated by the internal controller on the board (see Fig. 4).

The proposed controller is defined as (1), in order to compensate the dynamics of the robot.

$$W_{\mathrm{dj}}(t) = \mathrm{kp}_j \tilde{W}_j(t) + \mathrm{ki}_j \int_0^t \tilde{W}_j(t)\mathrm{d}t + \mathrm{kd}_j \frac{\mathrm{d}}{\mathrm{d}t} \tilde{W}_j(t) \tag{1}$$

where $\tilde{W}_j(t) = W_{\mathrm{dj}}(t) - W_{\mathrm{sj}}(t)$ with $j = 1, 2, 3, 4$ are the angular velocities error for each actuator;$\mathrm{kp}_j, k_j, \mathrm{kd}_j$ are positive gain that weigh the control errors.

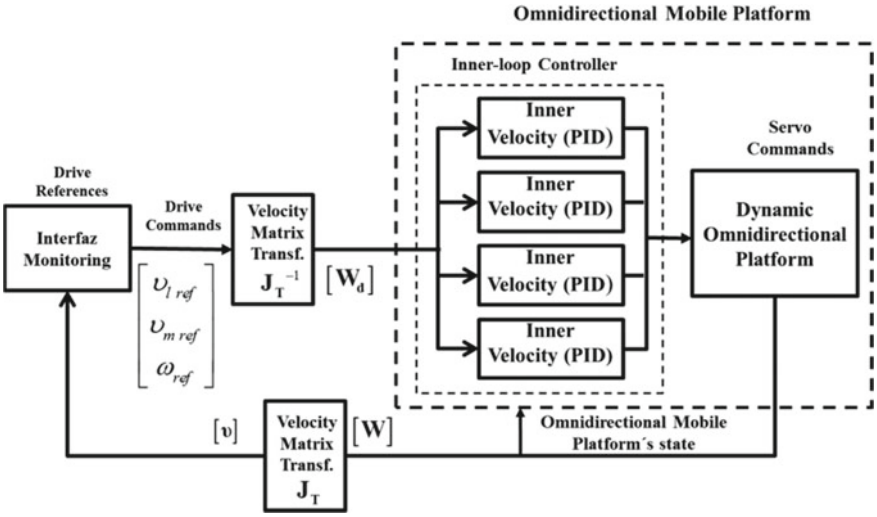


Fig. 4 Control scheme of the omnidirectional robot

Actuators: The actuators that make up the robot are DC motors at 11 V, which consist of an encoder incorporated in the axis of each actuator to measure the jumps and be transformed at angular velocities within the control task. The general system of the motors includes a driver that controls the velocity of each of them, i.e., receives a PWM signal to drive each motor.

Communication: The communication between the robot and the computer which generates the reference velocities to move to platform is wireless. For this the distance and amount of data to be transmitted, i.e., was established under the IEEE 802.15.4 standard that corresponds to Zigbee wireless networks.

3.1 Interface

The HMI was implemented in MATLAB software, since it presents great advantages when performing mathematical calculations to later apply a control for mobile platforms [12] it has the options to operate automatically the robot by the implementation of an advanced control or can be tele operated.

Automatic mode: Figure 5 in this mode, communication is established between the proposed control systems that generates the appropriate reference velocities $\boldsymbol{v_{ref}}$ to autonomously command the robot, for the case study these velocities will be generated by the operator to validate the operation of the prototype. The computer system receive these velocities and with the (6) will obtain the desired angular velocities for the robot $\boldsymbol{W_d}$.

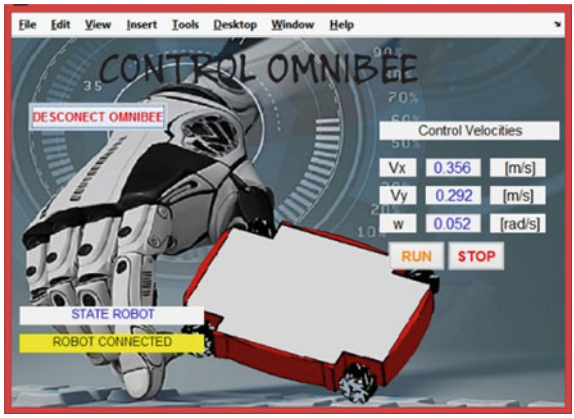


Fig. 5 Interface of automatic control

The angular velocities measured in the same way will be transmitted from the robot to the computer system thus obtaining a closed loop of the system to implement a control. In MATLAB, using (4), it determines the linear velocities υ_l, υ_m, and the angular ω measures in real time. At the end of the execution, the sent and received values from the omnidirectional robot will be displayed graphically to identify that it performs the requirements established by the operator at both reference velocities.

Tele-Operation Mode: Tele-Operation Mode: Fig. 6 This mode of operation of the omnidirectional robot is totally controlled by a user who generates commands by pressing the buttons of the HMI, which can be used to perform linear vertical linear/backward, horizontal right/left movements, rotate on your own axis right/left and diagonal displacements in 45°.

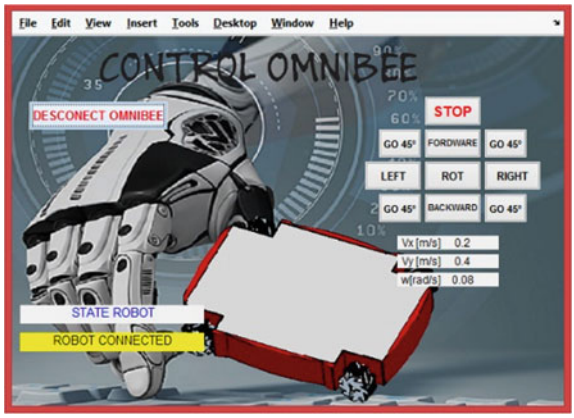


Fig. 6 HMI of the omnidirectional robot

4 Omnidirectional Robot Kinematics

Figure 7 shows the movement that the omnidirectional robot can perform on a horizontal plane.

Then the kinematic model of the omnidirectional robot is defined in the following way taking into account that the point of interest is the center of the robot (see Fig. 7),

$$\begin{bmatrix} \dot{p}_x \\ \dot{p}_y \\ \dot{p}_\psi \end{bmatrix} = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \upsilon_l \\ \upsilon_m \\ \omega \end{bmatrix} \tag{2}$$

Where $\dot{p}_x$, $\dot{p}_y$ and $\dot{p}_\psi$ represent the velocities with respect to the inertial reference frame. The equation system (2) can be written as,

$$\dot{\mathbf{p}}(t) = \mathbf{J}(\psi)\upsilon(t) \tag{3}$$

where $\mathbf{J}(\psi) \in \Re^{m \times n}$ with $m = n = 3$ is the Jacobian matrix that defines a linear mapping between the velocity vector of the omnidirectional robot $\upsilon \in \Re^m$ where $\boldsymbol{\upsilon} = \begin{bmatrix} \upsilon_1 & \upsilon_m & \omega \end{bmatrix}^T$ and the velocity vector $\dot{\mathbf{p}} \in \Re^m$ with $\dot{\mathbf{p}} = \begin{bmatrix} \dot{p}_x & \dot{p}_y & \dot{p}_\psi \end{bmatrix}^T$.

Figure 8 shows the configuration of the wheels of the omnidirectional mobile robot.

For find the velocities of the point of interest generated by the omnidirectional robot $\begin{bmatrix} \upsilon_1 & \upsilon_m & \omega \end{bmatrix}^T$, we take into account how the robot will be displaced with each direction of rotation of the omnidirectional wheels, thus obtaining the following system of equations:

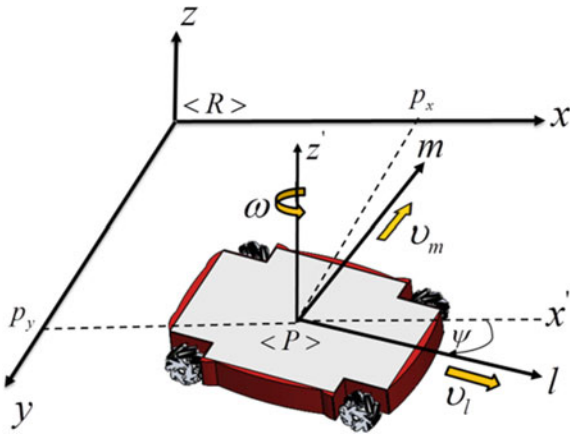


Fig. 7 Schematic of the omnidirectional mobile platform

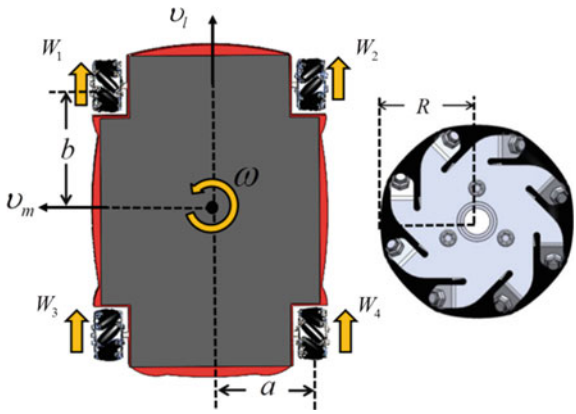


Fig. 8 Omnidirectional robot

$$\begin{bmatrix} \upsilon_1 \\ \upsilon_m \\ \omega \end{bmatrix} = \frac{R}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ \frac{-1}{a+b} & \frac{1}{a+b} & \frac{-1}{a+b} & \frac{1}{a+b} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \end{bmatrix} \tag{4}$$

The equation system (4) can be written as,

$$\boldsymbol{\upsilon}(t) = \mathbf{J}_T(a, b)\mathbf{W}(t) \tag{5}$$

where $\mathbf{W}$ is the vector of angular velocities of the wheels of the omnidirectional robot and $\mathrm{J}_T \in \mathbb{R}^{4\times3}$ is a transformation matrix, which was obtained from the direct kinematics of robot and $\boldsymbol{\upsilon}$ represents the mobility velocities of the omnidirectional robot.

The angular velocities $\mathbf{W_d}$ corresponding to each wheel of the platform are calculated based on the mobility velocities of the robot that can be generated by a controller or defined by the operator $\boldsymbol{\upsilon}_{\text{ref}}$, used (4) gives:

$$\mathbf{W}_d(t) = \mathbf{J}_T^{-1}(a, b)\boldsymbol{\upsilon}_{\text{ref}}(t) \tag{6}$$

Therefore the system of equations is defined as follows:

$$\begin{bmatrix} W_{d1} \\ W_{d2} \\ W_{d3} \\ W_{d4} \end{bmatrix} = \frac{1}{R} \begin{bmatrix} 1 & -1 & -(a+b) \\ 1 & 1 & (a+b) \\ 1 & 1 & -(a+b) \\ 1 & -1 & (a+b) \end{bmatrix} \begin{bmatrix} \upsilon_{l\,\text{ref}} \\ \upsilon_{m\,\text{ref}} \\ \omega_{\text{ref}} \end{bmatrix} \tag{7}$$

where $\boldsymbol{J}_{\boldsymbol{T}}^{-1} \in \mathbb{R}^{4x3}$ represent the inverse transformation matrix, which was obtained from the inverse kinematics of omnidirectional robot.

5 Results and Discussions

In order to validate the performance of the built robot, several experiments were carried out, the most representative results are described in this section. The experimental test was performed on the robot "OMNIBEE" name which was given to the omnidirectional platform built for the accomplishment of tasks that the operator proposes by means of controllers, this one supports two linear velocities and an angular velocity as signals of reference, see Fig. 9

For the tests of the robot, reference velocities are generated, and by means of (6) the velocities to be transmitted to the robot "OMNIBEE" are obtained. The robot transmits the angular velocities in a period of constant sampling, this will allow to verify that the robot executes in an ideal way each order that is applied to him in angular speed to his corresponding wheels. Figure 10 and 11 shown reference velocities applied to the robot and the velocities obtained.

Fig. 9 OMNIBEE omnidirectional mobile robot

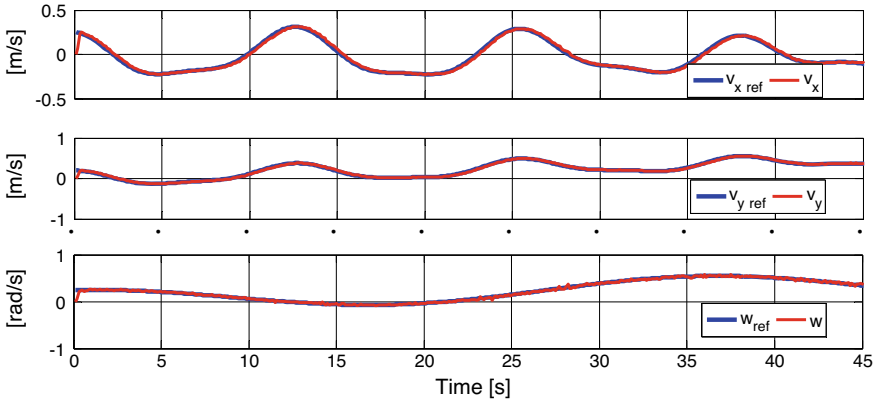


Fig. 10 Reference velocities of the omnidirectional robot and obtained velocities

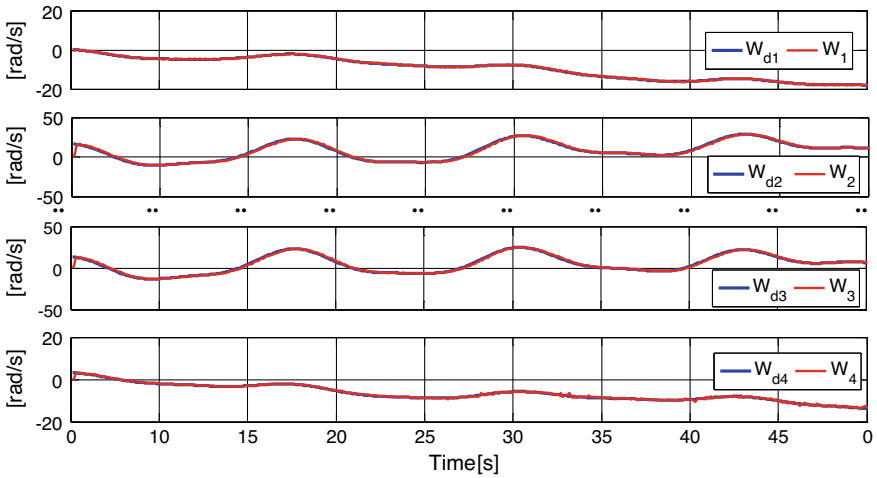


Fig. 11 Angular velocities sent and velocities received from the robot

For the second experiment, the OMNIBEE robot must complete a designated trajectory, using a controller. Figure 12 shows the strobe motion of the robot path in the plane X, Y, Z and Fig. 13 indicates the robot's maneuverability speeds to meet the trajectory.

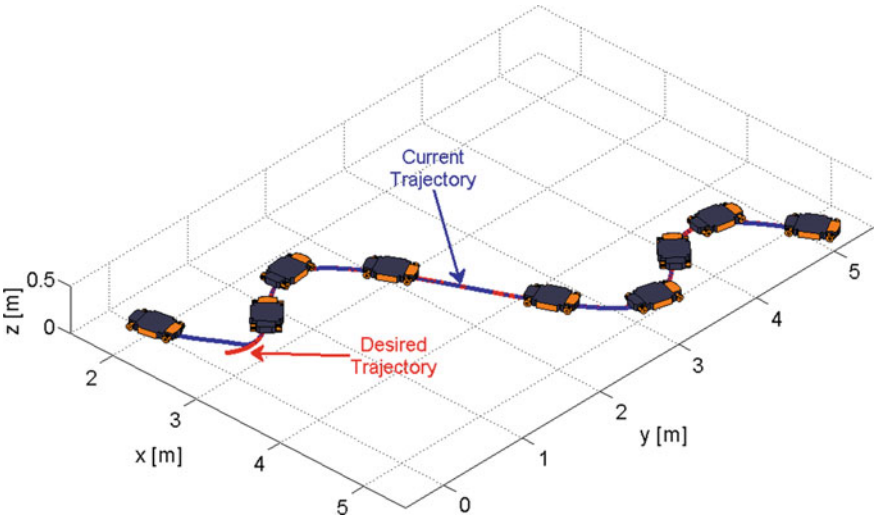


Fig. 12 Stroboscopic movement of the mobile robot

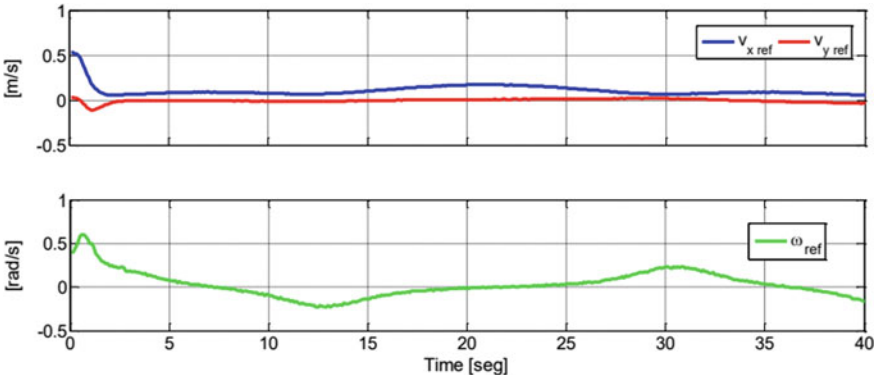


Fig. 13 Maniobrability velocity

6 Conclusions

The prototype to test new control methods and digital simulations helps greatly to understand concepts and model final products. Currently, there are digital tools and software that allow the design, simulation of parts and mechanisms that allow to verify the behavior of a mobile robot, with the possibility of analyzing its performance and efficiency of a finished model. In the project, we calculated the kinematic model that governs the mobile robot system, which allows a deeper analysis of our robot, and simulate closed loop control actions against angular velocity inputs, to analyze the trajectories of the robot on the plane X, Y, this allows to give solutions to problems of trajectory tracking, avoidance of obstacles, among other. For a future work can be added a manipulator to have automated controlled or tele operated mobile manipulator, to perform tasks of transporting objects or manipulation.

Acknowledgements The authors would like to thank the Universidad de las Fuerzas Armadas ESPE; Universidad Tecnológica Indoamérica; Universidad Nacional de Chimborazo; o Instituto de Automática de la Universidad Nacional de San Juan; SISAu Research Group, and the ARSI Research Group, for the support for the development of this work.

References

1. Morgan J, Thornton B, Peacock J, Hollingsworth K, Smith C, Oz M, Argenziano M (2005) Does robotic technology make minimally invasive cardiac surgery too expensive? A hospital cost analysis of robotic and conventional techniques. J Card Surg 20(3):246–251
2. Díaz C, Roa-Guerrero E (2015) Development of mobile robotics platform for identification of land mines antipersonal in different areas of Colombia. IEEE
3. Mondada F, Franzi E, Ienne P, Mobile robot miniaturisation: a tool for investigation in control algorithms. In: Lecture notes in control and information sciences, pp 501–513

4. James Mclurkin J (2017) Distributed algorithms for dispersion in indoor environments using a swarm of autonomous mobile robots. Citeseerx.ist.psu.edu
5. Jing Wang, On sign-board based inter-robot communication in distributed robotic systems. In: Proceedings of the 1994 IEEE international conference on robotics and automation
6. Stentz A (1994) Optimal and efficient path planning for partially-known environments. In: Proceedings of the 1994 IEEE international conference on robotics and automation, pp 3310–3317
7. Andaluz V, Ortiz J, Chicaiza F, Varela J, Espinosa E, Canseco P (2016) Adaptive control of the human-wheelchair system through brain signals. Intell Robot Appl 223–234
8. Zhang Y, Zhang L, Zhang X (2008) Mobile robot path planning base on the hybrid genetic algorithm in unknown environment. In: 2008 eighth international conference on intelligent systems design and applications
9. Shenoy P, Miller KJ, Crawford B, Rao RPN (2008) Online electromyographic control of a robotic prosthesis. In: IEEE transactions on biomedical engineering, vol 55, pp 1128–1135
10. Barraquand J, Kavraki L, Motwani R, Latombe J, Li T, Raghavan P (1996) A random sampling scheme for path planning. In: International symposium on robotics research, vol 6, pp 249–264
11. Walter J, Schulten K (1993) Implementation of self-organizing neural networks for visuo-motor control of an industrial robot. IEEE Trans Neural Netw 4(1):86–96
12. Andaluz V, Roberti F, Toibero J, Carelli R, Wagner B (2011) Adaptive dynamic path following control of an unicycle-like mobile robot. In: Intelligent robotics and applications, pp 563–574
13. Röhrig DH, Kirsch C, Künemund F (2010) Localization of an omnidirectional transport robot using IEEE 802.15.4a ranging and laser range finder. In: 2010 IEEE/RSJ international conference on intelligent robots and systems
14. Tlale, de Villiers M (2008) Kinematics and dynamics modelling of a mecanum wheeled mobile platform. In: 2008 15th international conference on mechatronics and machine vision in practice
15. Cooney WX, Bright G (2004) Visual dead-reckoning for motion control of a Mecanum wheeled mobile robot. Mechatronics 14(6):623–637
16. Kang W, Kim S, Chung J (2008) Development of omni-directional mobile robots with mecanum wheels assisting the disabled in a factory environment. In: International conference on control, automation and systems
17. Menzemer CC, Fei L, Srivatsan TS (1999) Design criteria for bolted connection elements in aluminum alloy 6061. ASME—J Mech Des 121:348–351

Flow Pattern Recognition Using Spectrogram of Flow Generated Sound with New Adaptive LBP Features

Soroosh Parsai and Majid Ahmadi

Abstract This paper represents the flow pattern classification utilizing the recorded sound of flows and their spectrograms synthesized images. The sound of four different flow patterns, stratified, churn, annular and slug flow, are recorded and converted to spectrogram for further analysis. The proposed method uses a new version of the local binary pattern (LBP) to extract robust features from the returned audio noise. We propose an adaptive threshold function based on Gaussian distribution's cumulative density function (CDF). The proposed algorithm performance is analyzed with both the 836 recorded sound of flow pattern and RWCP database to validate enhancements. The validation is done for two scenarios, one with the same noise signals for the training and test sets and one for different noises for each set. Furthermore, the new method is compared with three other methods (e.g., MFCC-HMM, SIF-SVM and MC-BDLBP). The comparisons show the new method's better performance with reduced complexity.

Keywords Flow pattern · Image processing · Local binary pattern · Spectrogram · Gaussian distribution

1 Introduction

Classifying the type of a 2-phase flow is an important practice in different industries, such as oil, gas, petrochemical and food. This specifically helps the process control system to change parameters to control the occurrence of them or prevent some unwanted damages that some of these patterns may create [1, 2]. When it comes to identifying these patterns, there are usually two main techniques in use. The first one is basically done by analyzing the parameters, e.g., finding the evidence of the existence

S. Parsai (✉) · M. Ahmadi
Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B3P4, Canada
e-mail: parsai@uwindsor.ca

M. Ahmadi
e-mail: ahmadi@uwindsor.ca

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_37

of a particular type of flow from the parameters such as velocity and pressure changes through the pipe. The second one more focus on the processing of the recorded sound of the flow. Understandably the parameter analysis method is not an automated technique and requires much more time for each instance. The latter method, however, falls into two major techniques: Putting the sound signal through the classification algorithms directly [23] or analyzing the image representation of the flow sound spectrogram via image processing techniques. Due to its cost-effectiveness, ease of implementation and vast variation of the image processing techniques, analyze of the spectrogram of the flow sound as an image seems more convenient method. The fact that this identification is made by image processing creates numerous possible methods for classification of flows. 2-phase flow classification is specified in two modes (1) a horizontal pipe and (2) a vertical pipe. In the cases that gravity is vertical to the pipe, phase separation occurs in a horizontal pipe. Some of the common 2-phase patterns in the industry are bubbly, Annular, plug, churn, stratified and slug flow [2].

The bubbly flow is defined as when a continuous stream of liquid exists with bubbles scattered through it. In the annular, liquid exists around the pipe's crust, and gas can be found around the center. The bubbles are unified in the plug flow pattern, and bigger bubbles can be found close to the pipe's wall. In cases of an oscillatory flow, we are dealing with Churn flow. The smoothness of the horizontal contact in a stratified pattern completely stratifies the liquid and gas. Slug flow is the other form, which is seen in the crude oil and tar production plants, particularly in pipelines conveying crude oil or liquid gas [3].

SVM and neural network feature extraction methods are the techniques used for 2-phase flow classification [4, 11]. The authors of this study [5] measure the performance of the airlift pump at various submergence ratios and compare the results to Hewitt and Roberts' flow pattern map [6]. Flow classification is represented using textural features and SVM classifier in [7]. Furthermore, a fuzzy neural network is proposed to identify the flow pattern of bubbly, slug and plug flow in [8]. However, according to the existing literature, flow characterization is facing some obstacles. Because these characterizations are dimensionally sensitive, they can only be used with the specific parameters of their research. Creating a universal flow pattern map for various liquids and pipe settings is a difficult undertaking. The fact that some flows may transit at a Weber number, and some might be justified by Reynolds number makes it almost impossible to come up with a universal dimensionless flow pattern map. Figure 1 depicts the occurrence of various 2-phase flow models in a horizontal pipe [9].

This paper presents flow type identification based on the audio signal classification to tackle the mentioned flow mapping problems. The hidden Markov model (HMM), which is a traditional speech-recognition technique and focuses on mel-frequency cepstral coefficients (MFCC), was used in the early research related to audio signal classification [10, 12]. Later on, other techniques like MPEG-7 audio features [13], matching pursuit features [14] and spectro-temporal signature [15] were developed as

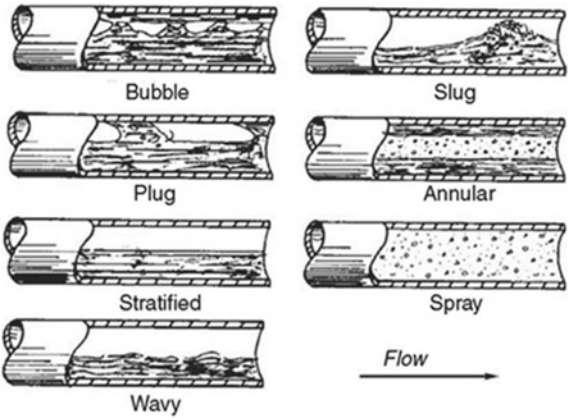


Fig. 1 Horizontal two-phase flow patterns [9]

the supplement for MFCC. Techniques such as mel-frequency cepstrum [16], short-time discrete Fourier transform (STFT) and Gabor spectrogram [17] are widely used in audio signal classifications due to their ability to work with spectrums.

The spectrogram of the audio signal provides helpful information in sound classification [24]. New approaches investigate spectrogram as a textured image [18–22, 25] and use image processing techniques for audio signal classification. The spectrogram is a composite image that illustrates the distribution of a signal's intensity in multiple frequencies at different time intervals. It is not a standard image. By recording the intensity differences between a pixel and its neighbors, the local binary pattern (LBP) approach retrieves the spectrogram's micro-patterns.

The LBP feature extractor has different forms, and all of these variations are commonly utilized in a variety of image processing applications [25, 26]. Typically, the LBP features obtained from the spectrogram have different applications. One application is related to audio signal classification [26]. The LBP features' drawback is their sensitivity to noise [27]. Different factors affect the reliability of the LBP features. These issues are mostly due to a lack of components in the LBP histogram. Processing patch–wise LBP is one of the problems that result in fewer components in each patch's histogram. In addition, the occurring frequencies patterns are vastly different, so these patterns cannot be precisely evaluated. The LBP feature extraction with an adaptive threshold is proposed to deal with such problems. The proposed method is used to classify and identify four flow patterns from sound signals. Furthermore, the new approach is compared with three works in [18, 19, 26] using the RWCP database.

The rest of the paper is organized as follows:

Section 2 briefly explains the experiment's setup for recording the flow's sound in the pipeline. The evaluation of the proposed method is discussed in Sects. 3. Section 4 presents the simulation results. Finally, in Sect. 5, the conclusion is represented.

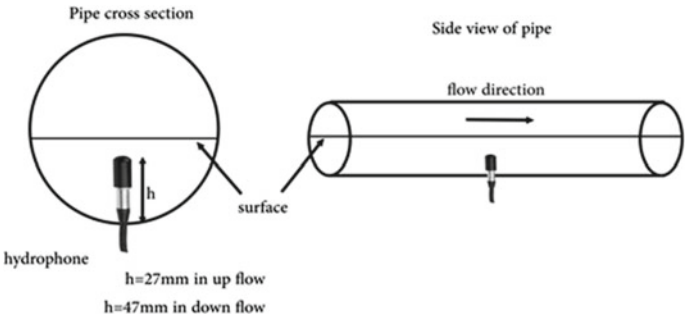


Fig. 2 Position of each hydrophone in the pipe

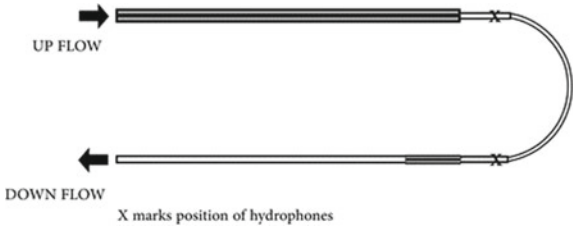


Fig. 3 Flow loop and the position of the two hydrophones (X marks the position of hydrophones)

2 Experiment's Setup

Two hydrophones are used in the setup, and both of them are connected by an amplifier to a signal analysis workstation. The position of the hydrophones is shown in Fig. 2. The system's input is the signals with a frequency range of 0.1–10 kHz. The diameter of the pipe's upward side and downward side is 5 inches and 3 inches, respectively. Figure 3 illustrates the schematic of the pipe with the position of two hydrophones.

3 Proposed LBP with Gaussian Reweighting Thresholding

3.1 Texture Analysis

In most applications, time–frequency analyses are used to create the spectrogram of the signals. The Gammatone filters and short-time Fourier transform are two instances of the existing approaches [18, 19]. The main difference between these two specific methods is in the spacing of the frequency positions on the axis. In gammatone-like spectrograms, frequencies have equal distances from each other and are scattered equally on the equivalent rectangular bandwidth (ERB). These properties make the gammatone approach a closer approximation of the human

sensory mechanism compared to the linear spaces of the frequency points on the STFT spectrogram. Time–frequency representations are suitable options to provide rich information about the texture. Of the main spectrograms, the Gammatone-like spectrogram (LGS) logarithm visually contains more texture information, making the LGS a suitable candidate.

For conducting the analysis, a collection of 50 Gammatone filters for the spectrogram are used. Frequency points on the ERB scale spaced in the interval of 100 and 10,000 Hz [18, 19]. The spectrogram of Gammatone is shown by $S(f, t)$, in which f represents the frequency, and t shows the time. For the improvement of the low-power instances, the following formula is used. This helps to gain more details, as illustrated in [26]:

$$G(f, t) = \log(S(f, t)) \tag{1}$$

The logarithm of the Gammatone-like spectrogram could be obtained by the following formula [26]:

$$I(f, t) = \frac{G(f, t) - \min\limits_{f,t} G(f, t)}{\max\limits_{f,t} G(f, t) - \min\limits_{f,t} G(f, t)} \tag{2}$$

3.2 Traditional LBP Features

Many feature extraction approaches have been proposed to analyze the image. Some of these approaches are SIFT [28], HOG [29] and LBP [30]. LBP is preferable among these feature extraction algorithms due to its easy implementation and effective results. It is crucial to consider the spectrogram as a synthesized image and acknowledge the many differences that exist between a spectrogram and a standard image. The following are some of the visual distinctions between a spectrogram and a standard image: [26]:

Smoothness: the first difference is that a standard image has a smooth dispersion, and the intensities change smoothly; on the other hand, in the spectrogram, the neighboring pixels could be remarkably different. The reason is that these pixels represent the power distribution of a sound signal, and they could be very different throughout the time–frequency instances.

Translation, scaling and rotation: The location and scales could vary for different shots in standard images. However, translation may only appear along the time axis in the spectrum. Hence, feature extraction methods such as SIFT, which extracts scale-invariant features, are not useful for spectrograms.

Micro-structure: Edges, spots and corners are usually considered essential details in natural images, but these micro-structures may not appear in the spectrogram. Therefore, edge-based feature extraction approaches such as HOG may not be appropriate for frequency analysis.

This paper considers LBP features the best way to classify flow patterns due to the following advantages. (1) The LBP feature works with signs of intensities which makes it resistant to spectrogram image monotonic changes. (2) In addition to edges and corners, LBP can extract other micro-structures which are not available to HOG. These advantages make the LBP a fit choice for the spectrogram feature extraction.

Conventional LBP algorithms use a fixed zero threshold [30]. The LBP code first runs on all the pixels. A pixel's intensity is denoted by i. The i_p determines the intensity of the p_{th} neighboring pixel, where $p = 1,2\ldots P$ and P is the number of neighbors. Here, R defines the distance between i_p and i. Considering mentioned notations, $\mathrm{LBP}_{p,R}$ encodes the pixel difference $z_p = i_p - i$ between i_p and i with R as the distance. Encoding every LBP bit with the following:

$$b_p = \begin{cases} 1 \text{ if } D_p \geq 0 \\ 0 \text{ if } D_p < 0 \end{cases} \tag{3}$$

The LBP code is formed as $\overrightarrow{b_p b_{p-1} \ldots b}$ from the mentioned LBP bits. After that, the $\mathrm{LBP}_{p,R}$, histogram of LBP codes with the dimension of 2^p is used as features vectors. It is evident that using more neighbors leads to more information on the image, increasing the dimension.

3.3 Proposed Method

Illumination variation is an important factor that affects the pixel intensity changes of the standard images. Thus, the pixel difference z_p also has a significant variation. It can be concluded from Eq. (3) that only the sign of the pixel variation is used by the LBP, and the amplitude is being dismissed. The amplitude of the pixel difference contains meaningful data describing the spectrogram. Therefore, both the direction and the amplitude of the pixel variation are used in [26] when extracting the patterns from the spectrogram. In conventional LBP features, a fixed zero threshold is proposed in [27], making it sensitive to noise. The noise can affect the system by changing the pattern bits for slight differences. A multi-channel LBP (MCLBP) feature to obtain the sign and amplitude of the pixel differences quantizes the pixel differences using multiple thresholds is proposed in [26]. One way to deal with this issue is to set the threshold of the pixel differences at T_i instead of thresholding at 0 [26, 33].

$$b_p^i = \begin{cases} 1 \text{ if } D_p \geq T_i \\ 0 \text{ if } D_p < T_i \end{cases} \tag{4}$$

where i indicates the channel number and b_p^i is the LBP's Pth bit in i. For each channel, a fixed user-defined threshold is used in this method. This paper proposes a new adaptive threshold function. This function is based on the local and global spectrogram information. This method does not make a different image band, but the proposed threshold changes in each frame. The proposed threshold function is based on a cumulative density function (CDF) of Gaussian distribution function to have both fast and robust LBP features, as below:

$$f_i = {}^1\!/_2\left[1 + \text{erf}\left(\frac{\sigma_i}{\sigma\sqrt{2}\sqrt{|\mu - \mu_i|}}\right)\right], i = 1, \ldots, M \tag{5}$$

where M is the number of spectrogram frames, σ and μ are standard deviation and global mean, and μ_i and σ_i are the mean and standard deviation in a frame of the image. The error function is described by erf(x) as bellow:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{6}$$

The proposed function is plotted in Fig. 4 for three different values of μ and σ.

Finally, each encoding LBP is obtained as below in adaptive threshold LBP (AT-LBP):

$$b_p^= \begin{cases} 1 \text{ if } D_p \geq f_i \\ 0 \text{ if } D_p < f_i \end{cases} \tag{7}$$

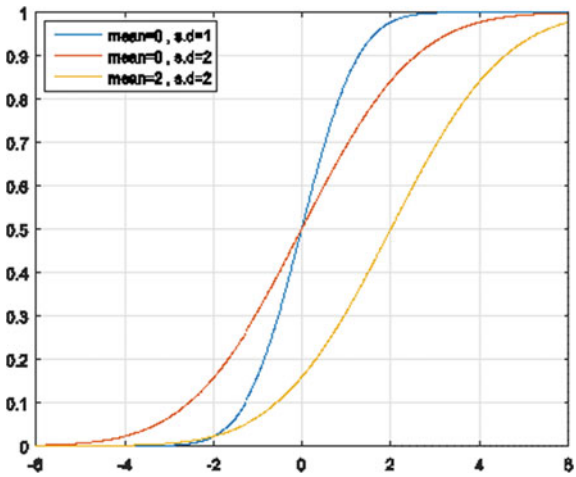


Fig. 4 The cumulative density function of Gaussian distribution with three different mean and standard deviation (s.d)

Noise properties have a determining role in the formation of the threshold function. When image noise increases, the threshold value will be increased and vice versa, leading to more robust LBP features.

4 Simulation Results

This section compares the flow pattern identification through spectrogram of recorded sound signals with the proposed method and conventional LBP. The proposed method is investigated for four flow patterns; Stratified, churn, annular and slug. In addition to that classification on the RWCP database is conducted for further validation of the proposed method and comparing the results to three other methods: MFCC-HMM [19], Spectrogram Image Feature (SIF) [18] and MC-BDLBP [26]. The RWCP database [31] is an appropriate benchmark in validating the performance of the classification algorithms in many research. The database consists of crash sounds of plastic, wood and other sounds like coins, bells, saw, etc. To study the effectiveness of the proposed method, the MFCC-SVM and SIF-SVM are applied by using the proposed AT-BLP features instead of MCC-HMM and SIF. The experiment is performed under four noise levels: 20, 10, 0 and −5 dB using the NOISEX92 database [32]. The applied linear SVM has a cost parameter set to 40 for classification. In the first experiment for the classification of pipeline flow, all the signals are sampled in 16 bits at 50 kHz using two hydrophones. The samples are collected from four classes of flow pattern types consisting of 836 samples in total. The number of samples for each class is shown in Table 1. In this experiment, 30% of samples are used for training and added noise is the same for training and test samples. As shown in Table 2, the proposed method has 98.7% correct classification for clean samples, whereas the traditional LBP represents 96%, and the new method shows much better performance when noises are added.

Table 1 Number of the samples for each flow pattern class

Number of samples	Class
184	Stratified
265	Churn
241	Annular
210	Slug

Table 2 Recognition rate of proposed method at different noise levels (%)

Method	Clean	20 dB	10 dB	0 dB	−5 dB
LBP	96.12	95.88	95.13	91.37	82.57
AT-LBP (proposed)	98.7	98.59	98.22	95.81	89.87

The RWCP contains 9722 sound events files with 105 distinct sets. The same partitioning of RWCP as in [18, 19, 26] is applied. A total of 50 audio signal stets are chosen from the RWCP database. The total 2500 and 1500 audio signals are considered for training and testing purposes. As in [26], two scenarios for adding noise are employed. In the first scenario, the SNRs are uniformly sampled from [−5 25] with the same type of noise in both training and test samples. The recognition rates are represented in Table 3. As shown in Table 3, the proposed method shows a better classification rate compared to the state of arts in [18, 19]. However, the method in [26] has better performance in the RWCP database at the cost of more complexity.

Table 4 illustrates the classification rates for the second scenario where the testing noise type is different from the training. The different noise types in the training and test samples cause a reduction in the recognition rate in this case. In this case, the proposed approach shows a better classification rate in most noise types than the MFCC-HMM and SIF-SVM. As shown in Table 3. The results are as par with MC-BDLBP while maintaining lower complexity.

5 Conclusions

This paper proposes a new LBP feature based on a new cumulative density function of Gaussian distribution thresholding. The algorithm is an enhancement of the conventional LBP algorithm. A comparison is made between the new algorithm and two other feature extraction techniques (e.g., MFCC-HMM and Spectrogram Image Feature (SIF)) through Mont Carlo simulations. Simulation results have shown the better performance of the proposed method by providing a better classification rate and less complexity.

Table 3 Percentage of the Classification rates in comparison with different methods under different noise sources on the RWCP Database in the first scenario

Method	Clean	Speaking noise				Destructive noise				Plant noise				Jet noise			
		20 dB	10 dB	0 dB	−5 dB	20 dB	10 dB	0 dB	−5 dB	20 dB	10 dB	0 dB	−5 dB	20 dB	10 dB	0 dB	−5 dB
MFCC-HMM	97.52	97.59	96.52	90.35	79.64	97.64	97.39	86.23	64.35	97.61	96.29	82.83	63.09	97.73	96.15	90.52	78.49
SIF-SVM	96.15	96.15	95.32	85.05	57.12	96.07	95.85	85.55	60.01	96.16	95.57	80.35	44.96	96.20	95.45	77.69	43.25
MC-BDLBP	99.80	99.71	99.79	99.65	99.53	99.80	99.71	99.03	96.85	99.80	99.60	97.63	95.20	99.87	99.77	99.61	99.45
AT-LBP (proposed)	99.32	99.11	98.86	98.24	97.89	99.12	98.92	98.64	96.26	99.77	99.56	97.24	94.87	99.76	99.59	99.43	99.28

Table 4 Percentage of the Classification rates in comparison with different methods under different noise sources on the RWCP database in the second scenario

Method	Clean	Speaking noise				Destructive noise				Plant noise				Jet noise			
		20 dB	10 dB	0 dB	−5 dB	20 dB	10 dB	0 dB	−5 dB	20 dB	10 dB	0 dB	−5 dB	20 dB	10 dB	0 dB	−5 dB
MFCC-HMM	82.71	82.36	82.21	63.84	46.33	97.64	97.39	86.23	64.35	97.61	96.29	82.83	63.09	97.73	96.15	90.52	78.49
SIF-SVM	67.59	67.91	68.78	49.12	25.94	96.07	95.85	85.55	60.01	96.16	95.57	80.35	44.96	96.20	95.45	77.69	43.25
MC-BDLBP	96.29	96.15	95.03	91.56	82.97	99.80	99.71	99.03	96.85	99.80	99.60	97.63	95.20	99.87	99.77	99.61	99.45
AT-LBP (proposed)	96.12	95.93	94.86	91.14	82.79	99.12	98.92	98.64	96.26	99.77	99.56	97.24	94.87	99.76	99.59	99.43	99.28

References

1. Dinh TB, Kim BS, Choi T-S (1999) Application of image processing techniques to air/water two-phase flow. In Applications of digital image processing XXII, vol 3808. International Society for Optics and Photonics, pp 725–731
2. Van Hout R, Barnea D, Shemer L (2001) Evolution of statistical parameters of gas-liquid slug flow along vertical pipes. Int J Multiphase Flow 27(9):1579–1602; Chen J, Richard C, Bermudez JC, Honeine P (2014) Variants of non-negative least-mean-square algorithm and convergence analysis. IEEE Trans Sig Proc 62(15):3990–4005
3. Shanthi C, Pappa N (2017) An artificial intelligence-based improved classification of two-phase flow patterns with feature extracted from acquired images. ISA Trans 68:425–432
4. Shanthi C, Pappa N, Suganya JA (2013) Digital image processing based flow regime identification of gas/liquid two-phase flow. IFAC Proc 46(32):409–414
5. Hanafizadeh P, Ghanbarzadeh S, Saidi MH (2011) Visual technique for detection of gas-liquid two-phase flow regime in the airlift pump. J Petrol Sci Eng 75(3–4):327–335
6. Hewitt GF, Roberts D (1969) Studies of two-phase flow patterns by simultaneous x-ray and flast photography. In: Atomic energy research establishment. Harwell, England (United Kingdom)
7. Huang G, Ji H, Huang Z, Wang B, Li H (2011) Flow regime identification of mini-pipe gas-liquid two-phase flow based on textural feature series. In instrumentation and measurement technology conference (I2MTC), 2011 IEEE. IEEE, pp 1–4
8. Shi L (2007) Fuzzy recognition for gas-liquid two-phase flow pattern based on image processing. In: Control and automation, 2007. ICCA 2007. IEEE international conference on, 2007. IEEE, pp 1424–1427
9. Bothamley M (2013) Gas/liquid separators: quantifying separation performance-Part 1. Oil Gas Facilities 2(04):21–29
10. Cowling M, Sitte R (2003) Comparison of techniques for environmental sound recognition. Pattern Recogn Lett 24(15):2895–2907
11. Guo G, Li SZ (2003) Content-based audio classification and retrieval by support vector machines. IEEE Trans Neural Netw 14(1):209–215
12. Temko A, Nadeu C (2005) Classification of meeting-room acoustic events with support vector machines and variable-feature-set clustering. In: Acoustics, speech, and signal processing, 2005. Proceedings.(ICASSP'05). IEEE international conference on, 2005. vol 5. IEEE, pp. v/505–v/508
13. Muhammad G, Alghathbar K (2009) Environment recognition from audio using MPEG-7 features. In: Embedded and multimedia computing, 2009. EM-Com 2009. 4th international conference on, 2009. IEEE, pp 1–6
14. Chu S, Narayanan S, Kuo C-CJ (2009) Environmental sound recognition with time-frequency audio features. IEEE Trans Audio Speech Lang Process 17(6):1142–1158
15. Tran HD, Li H (2011) Sound event recognition with probabilistic distance SVMs. IEEE Trans Audio Speech Lang Process 19(6):1556–1568
16. Logan B (2000) Mel frequency cepstral coefficients for music modeling. ISMIR 270:1–11
17. Qian S, Chen D (1999) Joint time-frequency analysis. IEEE Signal Process Mag 16(2):52–67
18. Dennis J, Tran HD, Li H (2011) Spectrogram image feature for sound event classification in mismatched conditions. IEEE Signal Process Lett 18(2):130–133
19. Dennis J, Tran HD, Chng ES (2013) Image feature representation of the subband power distribution for robust sound event classification. IEEE Trans Audio Speech Lang Process 21(2):367–377
20. Deshpande H, Singh R, Nam U (2001) Classification of music signals in the visual domain. In Proceedings of the COST-G6 conference on digital audio effects, pp 1–4
21. Paraskevas I, Chilton E (2003) Audio classification using acoustic images for retrieval from multimedia databases. In: Video/image processing and multimedia communications, 2003. 4th EURASIP conference focused on, 2003, vol 1. IEEE, pp 187–192

22. Yu G, Slotine JJ (2009) Audio classification from time-frequency texture. In: Acoustics, speech and signal processing, 2009. ICASSP 2009. IEEE international conference on, 2009. IEEE, pp 1677–1680
23. Matsui T, Goto M, Vert J-P, Uchiyama Y (2011) Gradient-based musical feature extraction based on scale-invariant feature transform. In: Signal processing conference, 2011 19th European, 2011. IEEE, pp 724–728
24. Kim MJ, Kim H (2012) Audio-based objectionable content detection using discriminative transforms of time-frequency dynamics. IEEE Trans Multimedia 14(5):1390–1400
25. Costa YM, Oliveira L, Koerich AL, Gouyon F, Martins J (2012) Music genre classification using LBP textural features. Signal Process 92(11):2723–2737
26. Ren J, Jiang X, Yuan J, Magnenat-Thalmann N (2017) Sound-event classification using robust texture features for robot hearing. IEEE Trans Multimedia 19(3):447–458
27. Ren J, Jiang X, Yuan J (2013) Noise-resistant local binary pattern with an embedded error-correction mechanism. IEEE Trans Image Proc 22(10):4049–4060
28. Geng C, Jiang X (2011) Face recognition based on the multi-scale local image structures. Pattern Recogn 44(10–11):2565–2575
29. Satpathy A, Jiang X, Eng H-L (2014) Human detection by quadratic classification on the subspace of extended histogram of gradients. IEEE Trans Image Process 23(1):287–297
30. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
31. Nakamura S, Hiyane K, Asano F, Nishiura T, Yamada T (2002) Acoustical sound database in real environments for sound scene understanding and hands-free speech recognition
32. Varga A, Steeneken HJ (1993) Assessment for automatic speech recognition: II. NOISEX-92: a database and an experiment to study the effect of additive noise on speech recognition systems. Speech Commun 12(3):247–251
33. Safaei N, Smadi O, Safaei B, Masoud A (2021) Efficient road crack detection based on an adaptive pixel-level segmentation algorithm. Transp Res Rec 2675(9):370–381

Analysis of Academic Excellence Achievement of Millennial Graduates through Attainment of the Learning Outcomes

Shikha Maheshwari, Kusum Rajawat, and Vijay Singh Rathore

Abstract This paper discusses the importance teaching–learning process for quality assurance and as efforts of affiliated technical institutes in India for getting accredited. Various problems related to the long-lasting acquisition of knowledge, skills, and attitude by the millennials are discussed with their implementable solutions to overcome it. The methodology suggested for attaining the course outcomes as well as program outcomes are supported with the values received after processing results of the learner-cohort for three consecutive years. The paper also discusses the action taken to fill the gap in the curriculum, if any.

Keywords Course outcomes · Outcome-based education · Program outcomes · Outcome attainment · Teaching–learning methodology

1 Introduction

It is a widely known fact that for sophisticated industrial products and services to sustain and flourish in the competition of the global market of the twenty-first century, and India will need a huge number of well-qualified graduates equipped with knowledge from autonomous as well as affiliated engineering institutions [1].

In such a scenario, it can be effectively identified that good teaching leads to effective learning, which leads to the thorough and long-lasting acquisition of the knowledge, skills, and attitude the educator, or the institution has set out to impart [2].

It is now understood that the quality assurance and accreditation efforts of engineering institutions can be characterized by their "compliance" rather than looking

S. Maheshwari (✉)
Manipal University, Jaipur, India
e-mail: shikha.maheshwari@onlinemanipal.com

K. Rajawat
Shri Karni College, Jaipur, India

V. Singh Rathore
IIS University, Jaipur, India

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_38

to it as an "improvement" tool. However, intellectuals have identified few problems, such as the need of designing an experiential curriculum for developing learners' experiences, tuning of learning outcomes with that of employers' expectations, collaboration with industry for designing, development, and effective delivery of the course/program content, adaptation of effective ICT-based methodologies for enhancing learners' experiences [3] as well as educators' competence and set-up of centers of excellence for the institutions in most sought-out technologies.

For providing solutions to all these mentioned pitfalls, various institutions both autonomous as well as affiliated implemented the following processes up to an extent to improve the quality of teaching and learning:

1. **Focus on outcome-based education** for the creation of an up-to-date experimental curriculum systematically providing learners' experiences in synchronization with the employers' expectations to fulfill the society's requirements.
2. **Faculty enablement** in terms of pedagogical and engineering domain training is required, which is expected to initiate the interactive teaching–learning process in every classroom. Pedagogy is required to include modern teaching tools such as simulations, group presentations and discussions, extempore, and reviewing case studies.
3. **Leveraging technology** for facilitating the integration of in-classroom as well as outside-classroom learning for assessment of learning outcomes across the two venues, thus enabling learners to reflect on and share their experiences and outcomes. Digital technologies also enable educators to concentrate on learning activities enhanced by state-of-the-art content delivery [4].
4. **Foster stronger industry-institute collaborations** can be achieved by making **industry internships mandatory practice** for all engineering learners as industries have good opportunities to develop, produce, and deliver affordable and quality products/services to cater for the fundamental needs of a growing population. This step will induce the innovative and entrepreneurial nature of engineering graduates for economic prosperity and wealth in the country. Also, there can be a c**ollaboration with industry for research and teaching** by creating consortia for applied research and engineering, and there can be sabbatical rotation arrangements for educators to take up application-driven joint research collaboration, curriculum development, and continuous education with industry and vice versa for the industry personnel to join as educators at an institute, with some sort of seed-funding from the government as well. Such steps will surely build global competitive clusters of companies, institutes as well as research organizations.
5. **Encourage innovation and entrepreneurship** by establishing **incubation centers**/warehouses, with the assistance and mentorship of industry councils initially till the institutes can take off on their own. Focus areas can be joint research projects, industry consultancy, and IPR generation in creating innovation eco-systems both in specific disciplines

To achieve academic excellence, classroom attendance still plays a crucial role. With the advent of technology, educators may adopt a two-pronged strategy, in which one

may have to implement the 75% attendance criteria a compulsory in letter and spirit and later, approach makes the classroom teaching interesting and lively with a lot of simulations providing a visual supplement to reading and along with a lot of material, not available in textbooks. This also requires to be accompanied with an explanation of some of the difficult concepts with examples, of course, relevant but day-to-day life by the educators during their content delivery. Not only this, effective mentoring for understanding and sympathetic toward the learners' problems facilitates the generation of balanced questions banks covering the entire curriculum for preparing well for their examinations, soliciting mid-semester learners' feedback for hearing their concerns for planning appropriate action, and creating an environment, in which the educators feel the pride of belongingness to the institution, derives maximum output toward academic excellence [5].

It is widely accepted that accreditation initiatives taken by the National Board of Accreditation in India are intended to promote quality in Engineering Education to produce graduates for a long-term strategic impact on growth in industrial and service sectors of any country, according to the Washington Accord. An institution that has successfully achieved NBA accreditation carries the stamp of recognition for its quality assurance in educational activities. The process of accreditation is normally carried out through data collection under a given set of criteria in a self-assessment report to be filled up by each institute that wishes to undergo the same, which is then evaluated by the board. One such criterion is criteria 3 which evaluates the course-wise performance for each learner and helps in identifying the weaker learners and planning remedial actions for them [6, 7].

For this paper, the curriculum of few state universities is considered, and a dataset of an affiliated institute is used for the study purpose. The curriculum is comprising of mathematics, core electronics, and core courses concerning the computer science and engineering branch. For identifying the gaps, courses are mapped with the set of program outcomes as listed by the National Board of Accreditation with course classification under categories such as advanced mathematics, core electronics and core professional, and elective [8, 9]. In order to meet its educational goals, the institute has established an exhaustive set of measurable program educational objectives, PEOs, which for each department lists the characteristics of graduates. With this information, they will be well-equipped to obtain careers in their respective fields. PEOs are used to describe the career and professional accomplishments expected by the learners to be achieved in five years after graduation with respect to the program's constituencies and the institute's mission [10].

The 12 graduate attributes are designated as program outcomes representing the knowledge, skills, and attitudes of the learners expected to have at the end of their program in India [11].

Additionally, the departments have established a set of program specific outcomes, detailing what engineers graduating from a particular program should be able to accomplish.

All courses relate to elementary ideas, tools, and procedures, and all have an emphasis on practical applications that culminate in increased impact on many

aspects of the lives of people throughout the world. By allowing this, a strong connection is established between course outcomes and program outcomes, while at the same time providing learners the competencies, they will need to be well-rounded engineers. With this intention, the process has been used to identify the extent of compliance of the university curriculum for the attainment of the POs and PSOs with the proper mention of the identified curricular gaps [12].

Not only this, the significant relationship between the courses and its program outcomes (POs) and program specific outcomes (PSOs), if any, is established with the help of well-defined cognitive processes a course provides in the form of course outcomes statements, COs. The COs are the resultant knowledge or skills or attitude the learner acquires at the end of a course. Though evidence of COs being defined for every course is available, for our analysis, only core courses third-semester onwards of the computer science engineering program is considered, which comes out to be five courses, one per semester [13]. These COs are then mapped with POs as well as PSOs using the matrix, two out of the six selected courses as shown in Fig. 1 and Table 1.

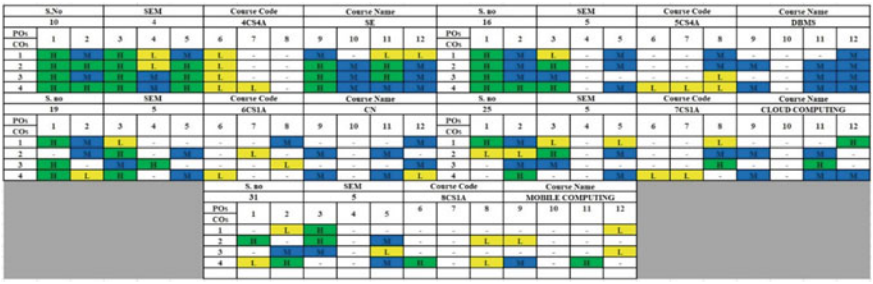


Fig. 1 Mapping of COs with POs

Table 1 Mapping of COs with PSOs

S.No.	Sem	Code	Course	COs	PSO1	PSO2
10	4	4CS4A	SE	CO1	L	M
				CO2	L	M
				CO3	M	M
				CO4	M	M
16	5	5CS4A	DBMS	CO1	H	M
				CO2	H	M
				CO3	H	M
				CO4	H	M
19	6	6CS1A	CN	CO1	-	
				CO2	-	-
				CO3	-	-
				CO4	M	M
25	7	7CS1A	Cloud Computing	CO1	-	-
				CO2	-	-
				CO3	-	L
				CO4	-	L
31	8	8CS1A	Mobile Computing	CO1	H	H
				CO2	H	M
				CO3	H	H
				CO4	H	H

2 Methodology for Attainment of Outcomes

The procedure of assessing compliance with the university curriculum is used to identify the level of conformity with the CO. This includes mainly lecturing, taking tutorials, presentations, experimental laboratory work, group tasks, handouts, social activity, FDP/conference/workshop, e-book/digital library/video lectures, hackathon, Swayam, TedX/MUN effectively to teach learners about a concerned course. Teachers impart an in-depth understanding of history, ideas, concepts, theories, and analogies to make the content easy to understand by lectures. Educators help the slow learners by solving a greater number of similar problems during the tutorial sessions; also, specific problems are solved during such sessions. Educators also provide course content in the form of presentations and videos which effectively communicate the working of actual engineering solutions, therefore, leading to life-long learning in the appropriate societal context. This is further assisted through a type of laboratory work in which theories are put to the test by conducting experiments and interpreting the findings. Tutorials and laboratory work are generally carried out as group task which effectively enables the showcasing of the engineering concepts understood by the learners. Handouts, giving a quick insight into the course, are provided to help the slow learners to face the examinations with confidence.

In addition to this, digital library resources such as NPTEL and YouTube, which are comprised of e-books and e-journals, are available where interested learners and instructors may read. Using these resources, visualized learning activities are also organized in small groups for the whole cohort. To provide the best teaching–learning resources, interactive Swayam courses, prepared by the best educators in the country, are also suggested to the learners which became very popular among them.

IBM has established a center of excellence lab at many institutes so that learners can learn the latest software and techniques which are developed by IBM like Tivoli. Hackathon and Smart India Hackathon were organized on the campuses of these technical institutes to promote IT and e-governance initiatives. TEDx talk and Model United Nations organized to discuss various topics by the eminent personalities and offered contestants to demonstrate their ability by researching, thinking critically, and public speaking [14].

To give a sense of social responsibility to learners, various social activities were organized on the campus which includes blood donation, cleanliness drives, and foremost offering a helping hand to the underprivileged section of society by providing them tuition for free after college hours.

Not only learners but educators are actively engaged in the participation for faculty development programs, workshops, and conferences to aid them in the dissemination of knowledge about the latest technology to their learners [15].

Attainment of program outcomes and program specific outcomes has also been recorded using curriculum, assessment, and evaluation as the assessment tools and processes.

3 Result and Analysis

The attainment of course outcomes of all courses concerning set attainment levels has been recorded for three consecutive years of the pre-pandemic era are as in Table 3.

A year-wise comparison of course outcomes then reviewed for each of these chosen courses and one such CO attainment is presented in the Fig. 2.

Applying the concepts of outcome-based education for calculating the attainment of program outcomes, an attainment target is set for each PO. The reason behind choosing only one course is to reduce the complexity of the analysis presented in this paper (Fig. 3) (Table 2).

Gap is then calculated using the attained values for each PO with respect to ideal and target value.

Here, it should be noted that only one course is considered for showing the calculations related to the gap identification to reduce the complexity of the topic and for its better understanding.

It was observed that the teaching–learning methodology suggested for attaining the course outcomes has successfully reduced the gap with the advancement in the year, which can be seen in the Fig. 4 (Fig. 5).

POs	Tool	Academic Assessment50%					Placement 20%					Beyond Curriculum 20%					Feedback 10%		
	Tools	Sessional Mid-Term Test Result	Final University Result	Student Project	Lab/Experiments	Industrial training	Core Sector	Mentoring	Soft Skills	Higher Studies	PSU/GATE	Technical Events	Conference/Workshops	Social Activities	E-Resources	Industrial visit	Course Exit	Student Exit	Alumni Exit

Fig. 2 Tools and processes used for assessment to measure the attainment of the program outcomes

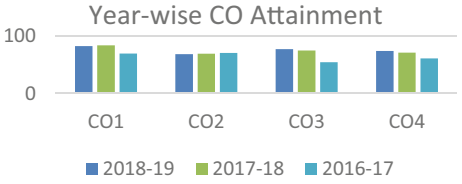


Fig. 3 Year-wise comparison of CO attainment for one course

Table 2 Achievement of PO attainment for one course

POs		POI	P02	P03	P04	P05	P06	P07	P08	POP	POlO	POll	P012
Target %		70.4	64.S	74.1	74.1	70.4	70.4	55.6	59.3	83.3	57.4	70.4	66.7
2016–17	4CS4A	29	29.6	34.1	31.2	31	31.8	24	23.8	34.5	24.8	28.4	28.6
2017–18	4CS4A	39.4	50.6	47.8	50.9	50.2	51	44.8	39.3	50.4	42.9	50.4	45.8
2018–19	4CS4A	49	59.4	55.2	56	55.4	56.1	56.5	48.7	56.5	45.6	59.3	55.9

PO Attainment (2016-17) for 4CS4A												
PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
Actual /Ideal Value	75	75	75	75	75	75	75	75	75	75	75	75
Target Value	60	60	60	60	60	60	60	60	60	60	60	60
Target Value%	80	80	80	80	80	80	80	80	80	80	80	80
Attained Value	29	29.6	34.1	31.2	31	31.8	24	23.8	34.5	24.8	28.4	28.6
Attained Value % wrt Ideal	38.67	39.47	45.47	41.6	41.33	42.4	32	31.73	46	33.07	37.87	38.13
Gap%	41.33	40.53	34.53	38.4	38.67	37.6	48	48.27	34	46.93	42.13	41.87

PO Attainment (2017-18) for 4CS4A												
PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
Actual /Ideal Value	75	75	75	75	75	75	75	75	75	75	75	75
Target Value	60	60	60	60	60	60	60	60	60	60	60	60
Target Value%	80	80	80	80	80	80	80	80	80	80	80	80
Attained Value	39.4	50.6	47.8	50.9	50.2	51	44.8	39.3	50.4	42.9	50.4	45.8
Attained Value % wrt Ideal	52.53	67.47	63.73	67.87	66.93	68	59.73	52.4	67.2	57.2	67.2	61.07
Gap%	27.47	12.53	16.27	12.13	13.07	12	20.27	27.6	12.8	22.8	12.8	18.93

PO Attainment (2018-19) for 4CS4A												
PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
Actual /Ideal Value	75	75	75	75	75	75	75	75	75	75	75	75
Target Value	60	60	60	60	60	60	60	60	60	60	60	60
Target Value%	80	80	80	80	80	80	80	80	80	80	80	80
Attained Value	49	59.4	55.2	56	55.4	56.1	56.5	48.7	56.5	45.6	59.3	55.9
Attained Value % wrt Ideal	65.33	79.2	73.6	74.67	73.87	74.8	75.33	64.93	75.33	60.8	79.07	74.53
Gap%	14.67	0.8	6.4	5.33	6.13	5.2	4.67	15.07	4.67	19.2	0.93	5.47

Fig. 4 Identification of gap in PO attainment

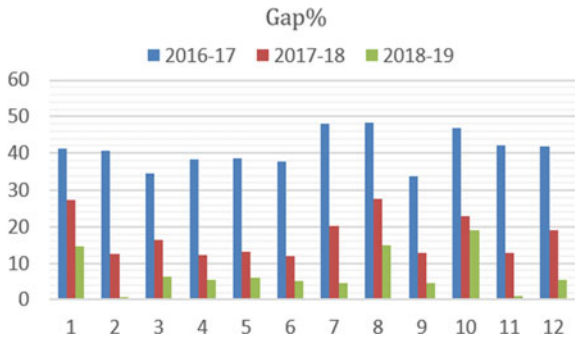


Fig. 5 Analysis of gap% in PO attainment year wise for one course

Similarly, PSOs attainment target was also set. Gap % also calculated for PSO attainment on the same lines as that of PO attainment, and it can be easily seen that this gap in PSO attainments is also reducing with the advancement of year (Fig. 6).

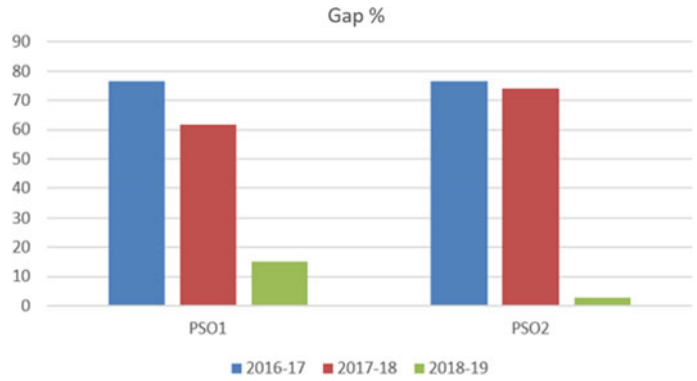


Fig. 6 Analysis of gap% in PSO attainment year wise for one course

4 Conclusion

The conclusion of this research paper is focusing on the actions proposed on the basis results of evaluation of each of the POs and PSOs. It is correctly said that curriculum and the teaching/learning processes are the foundation on which any program can be developed. Attainment of POs indicates that the job is well done. While identifying the gap for PO attainment, certain deficiencies were identified and proper measures then suggested by the expert educators to overcome it. Few deficiencies identified are listed under:

1. Learners should learn to use effective knowledge based on the latest research for problem-solving and project development.
2. Learners were unable to consider various ideas' feasibility and adhere to finances.
3. It is an industry requirement to have new tools and software, in order to bridge the gap between industry and academic.
4. Content not included in the syllabus should be part of a comprehensive curriculum connected to the public's health and safety, as well as the social needs.
5. Learners should approach their relationship with the environment and global consciousness.
6. An awareness of ethical and technical knowledge is part of every modern graduate's education.
7. There are some courses in the curriculum that do not properly train learners for their roles as an individual and a team member in diverse situations as well as varied collaborations.
8. The ability to communicate effectively is something that needs improvement for presentations to be effective. In addition, learners lack report writing and design documentation skills, which results in their being ineffective while making presentations.

The above-stated concerns could be addressed by including more workshops and other technical activities in the curriculum to increase learners' capacity to relate it to classroom lectures. Engaging videos were used to inspire the learners to apply up for online courses. The learners were instructed to look at the challenges and issues that are encountered in real life. Increase in home assignments having computational importance along with remedial classes for all such courses in which target is not achieved were conducted. Proper feasibility analysis and design and development of the product were suggested to learners to help them meet the industry needs in a multi-disciplinary environment. Design solutions for complex engineering challenges and coding contests offered to learners in addition to programs with these focus points. When learners are inspired to use their engineering skills for social impact, they have the knowledge required to tackle numerous societal issues. Multiple additional activities were also undertaken on professional engineering solutions for the environment in order to strengthen their leadership abilities while also giving them the opportunity to work in teams. Sustainable development knowledge is spread within a community of learners. A motivational lecture series was established to promote self-actualization, to teach people about ethical principles, and to show the need of keeping a professional reputation. Additional to this, learners wanted to collaborate together on their technological projects. Learners learned diverse communication, technical, and presentation skills while also acquiring personality development skills. Also, expert talks were conducted to enhance aptitude, qualitative skills of the learners. Additional classes were conducted for writing effective reports and design documents. Similarly, learners were encouraged to participate in entrepreneurship and startups programs.

References

1. https://www.nbaind.org/files/general-manual-of-accreditation.pdf
2. https://www.aspiringminds.com/research-reports/national-employability-report-for-engineers-2019/
3. Venkatesh P (2013) Accreditation as a benchmark of quality sustenance in higher education: a review of quality assurance in Indian university system. Nitte Manage Rev 7(1):59–67
4. https://www.nbaind.org/files/obe-and-nba-accreditation.pdf
5. Stefan BF (2006) Globalization, curricula reform and the consequences for engineers working in an international company. Eur J Eng Educ 31(3):261–272
6. Chowdhury H, Alam F, Biswas SK, Islam MT, Islam AS (2013) Quality assurance and accreditation of engineering education in Bangladesh. Procedia Eng 1(56):864–869
7. Killen R (2000) Outcomes-based education: principles and possibilities. Unpublished manuscript, University of Newcastle, faculty of education, Aug 2000
8. Al-Yahya SA, Abdel-Halim MA (2012) A successful experience of ABET accreditation of an electrical engineering program. IEEE Trans Educ 56(2):165–173
9. Shearman R, Seddon D (2010) Challenges for academic accreditation: the UK experience. Eur J Eng Educ 35(4):469–477
10. Ezeldin AS (2013) International accreditation for engineering programs: mission, learning objectives and outcomes. Procedia Soc Behav Sci 22(102):267–275

11. Min KJ, Jackman JK, Gemmill DD (2013) Assessment and evaluation of objectives and outcomes for continuous improvement of an industrial engineering program. Int J Eng Educ 29(2):520
12. Turhan C, Sengul G, Koyuncu MA (2015) Comprehensive assessment plan for accreditation in engineering education: a case study in Turkey. Int J Eng Educ Jan 1 31(2015):1270–81
13. Augusti G (2012) Quality and accreditation in engineering education in Europe. In: International symposium on innovation and quality in engineering education. Valladolid (Spain), Apr 2012, pp 26–28
14. https://www.nbaind.org/files/co-po-mapping-and-attainment-of-cos-and-pos.pdf
15. Rajak A, Shrivastava AK, Bhardwaj S, Tripathi AK (2019) Assessment and attainment of program educational objectives for post graduate courses. Int J Modern Educ Comput Sci 1(2):26–32

Business Intelligence in Strategic Business Decision Making in Times of COVID-19: A Systematic Review of the Literature

Alexis Carbajal-Torres, Joseph Ninaquispe-Florez, and Michael Cabanillas-Carbonell

Abstract Today, as a result of the pandemic we are going through due to the infectious disease COVID-19, many small, medium and large companies worldwide have had difficulties to stay in the market, because they have not made appropriate strategic decisions causing most of them to close. It is therefore important that companies adapt business intelligence and note the effect it has on them. The present investigation is a systematic review of the literature, compiling 55 articles from the following databases; Springer Link, IEEE Xplore, Wiley and Scopus that according to our inclusion and exclusion criteria 26 articles were systematized showing how business intelligence adapts and positively influences companies.

Keywords COVID-19 · Business intelligence · Strategic decisions · Companies · Systematic review

1 Introduction

The COVID-19 pandemic has notoriously affected small and large companies, being the trigger for many of them that led them to bankruptcy.

For this reason, it is necessary to combine knowledge from different areas to have an interdisciplinary vision and face the problem; having knowledge of business intelligence establishes bridges between the objectives in search of solutions to strengthen the business model [1].

For this purpose, we intend to know the adoption of Business Intelligence in these companies, followed by the effect it caused. BI being of utmost importance

A. Carbajal-Torres · J. Ninaquispe-Florez
Universidad Autónoma del Perú, Lima, Perú
e-mail: acarbajalt@autonoma.edu.pe

J. Ninaquispe-Florez
e-mail: jninaquispe@autonoma.edu.pe

M. Cabanillas-Carbonell (✉)
Universidad Privada del Norte, Lima, Perú
e-mail: mcabanillas@ieee.org

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_39

for decision making, which can be applied in such a way that it is very useful by identifying patterns. In this case applied to the pandemic, improving management response [2].

For this reason, a systematic review of the scientific literature will be carried out in order to generate structured data for decision making [3].

2 Methodology

2.1 Type of Study

This research is a systematic review of the scientific literature, which allows the collection of relevant evidence on a given topic, which, according to the selection criteria, responds to the research questions formulated.

2.2 Research Questions

The proposed research questions are as follows:

RQ1. What Business Intelligence strategies have been used in COVID-19 times for decision making in the business sector?

RQ2. What have been the benefits of Business Intelligence in times of COVID-19 for decision making in the business sector?

RQ3. Which countries have developed the most research on the topic of Business Intelligence in strategic decision making in times of COVID-19?

RQ4. Which technologies other than Business Intelligence support strategic business decision making in times of COVID-19?

2.3 Search Strategies

A search for articles was then carried out in the main databases such as Springer Link, IEEE Xplore, Wiley and Scopus. Of which 55 scientific articles were collected. In conducting the search to answer our research questions, the following keywords were taken into consideration: "Business Intelligence AND (decision making OR strategic business decision making) AND COVID-19" (Fig. 1).

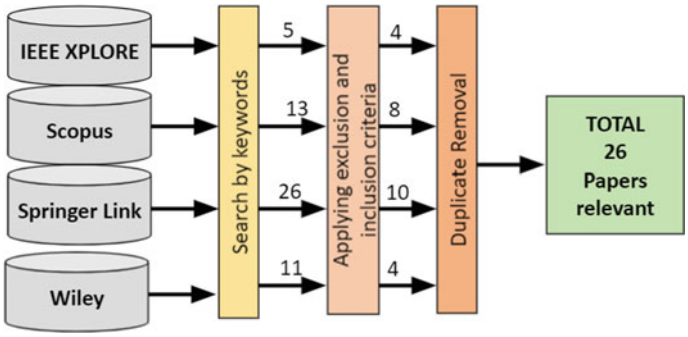


Fig. 1 Inclusion and exclusion stage

Table 1 Criteria inclusion and exclusion

	Criterion	
Inclusion	I01	Articles related to Business Intelligence in times of COVID-19
	I02	Articles related to decision making in times of COVID-19
	I03	Articles published in the years 2020–2021
Exclusion	E01	Articles not related to Business Intelligence in times of COVID-19
	E02	Articles related to COVID-19 but not focused on Business Intelligence
	E03	Articles not published in the years 2020–2021

2.4 Inclusion and Exclusion Criteria

The following inclusion and exclusion criteria were used for data analysis (see Table 1).

Fifty-five articles collected from the databases were studied; 15 articles that did not meet the inclusion criteria and 13 that did not contribute to answering the research questions were excluded. In the end, 26 articles were available for meta-analysis (Fig. 2).

3 Results

Figure 3 shows the number of articles analyzed by database.

Figure 4 shows the number of articles analyzed by country.

Figure 5 shows the number of strategies applied in business intelligence according to the articles analyzed.

Figure 6 shows the number of technologies mostly applied in business intelligence according to the articles researched.

Figure 7 shows the number of items per benefit.

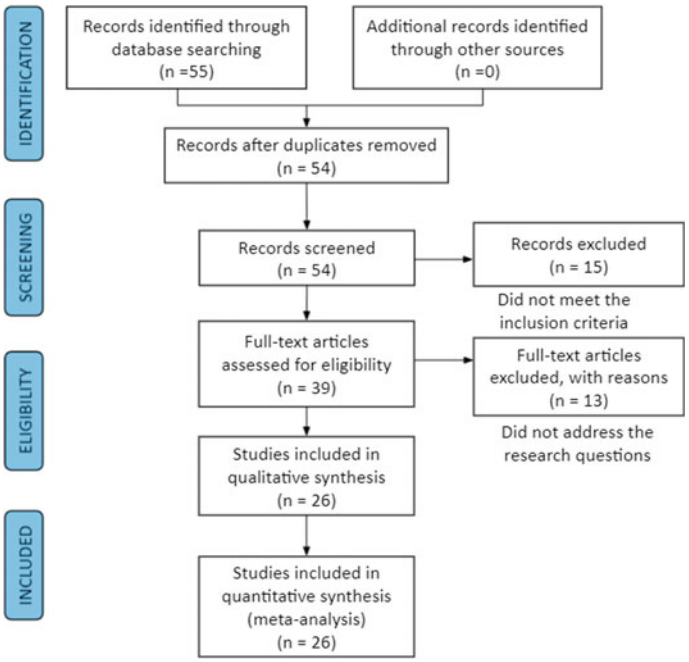


Fig. 2 Prisma Diagram Methodology

Fig. 3 Articles per database

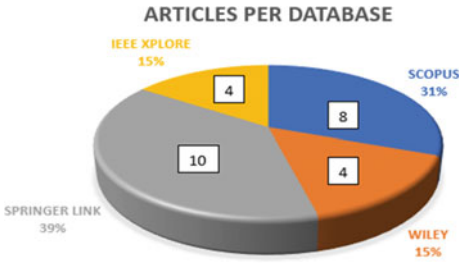


Table 2 shows the strategies of the strategic decisions according to the results found.

Table 3 shows the strategic decision technologies according to the results found.

Table 4 shows the benefits of the strategic decisions according to the results found.

4 Discussions

RQ1. What Business Intelligence strategies have been used in COVID-19 times for decision making in the business sector?

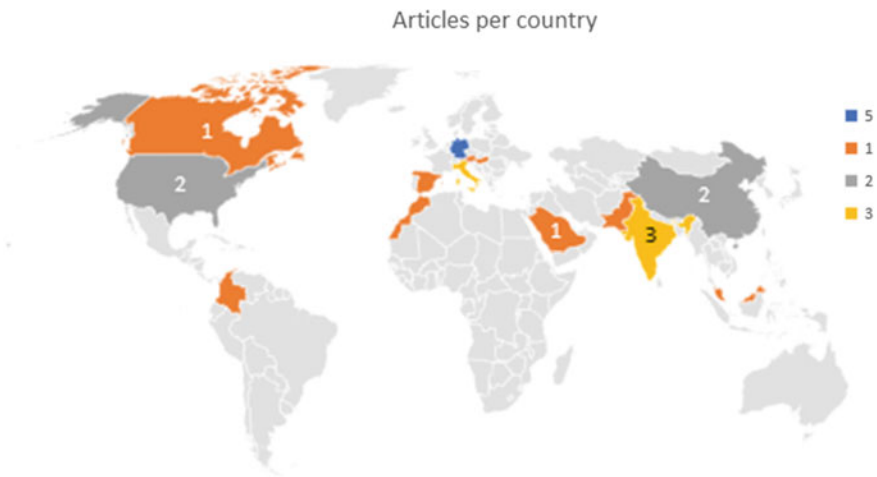


Fig. 4 Articles by country

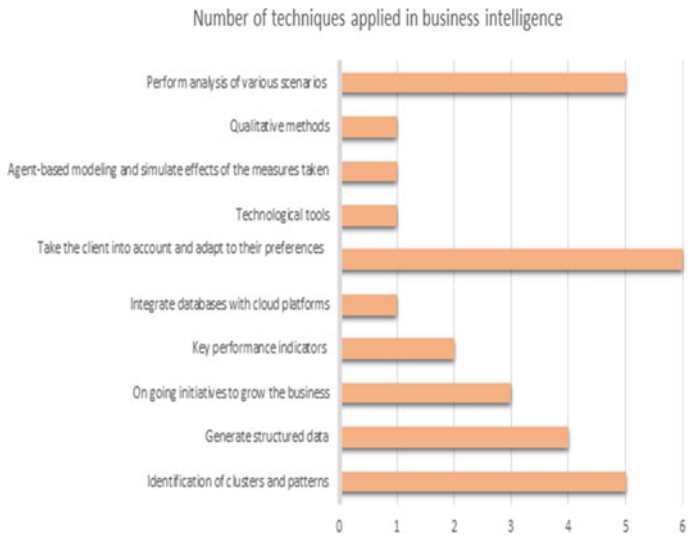


Fig. 5 Number of strategies applied in business intelligence per item

According to Table 2 and Fig. 5, it can be observed that the articles related to the topic have strategies such as generating structured data, key indicators, qualitative methods, performing analysis of various scenarios. This result indicates that these techniques are being used for strategic business decision making in times of COVID-19.

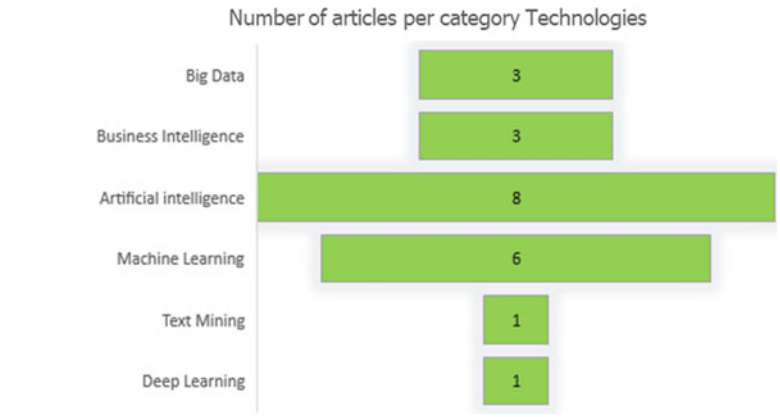
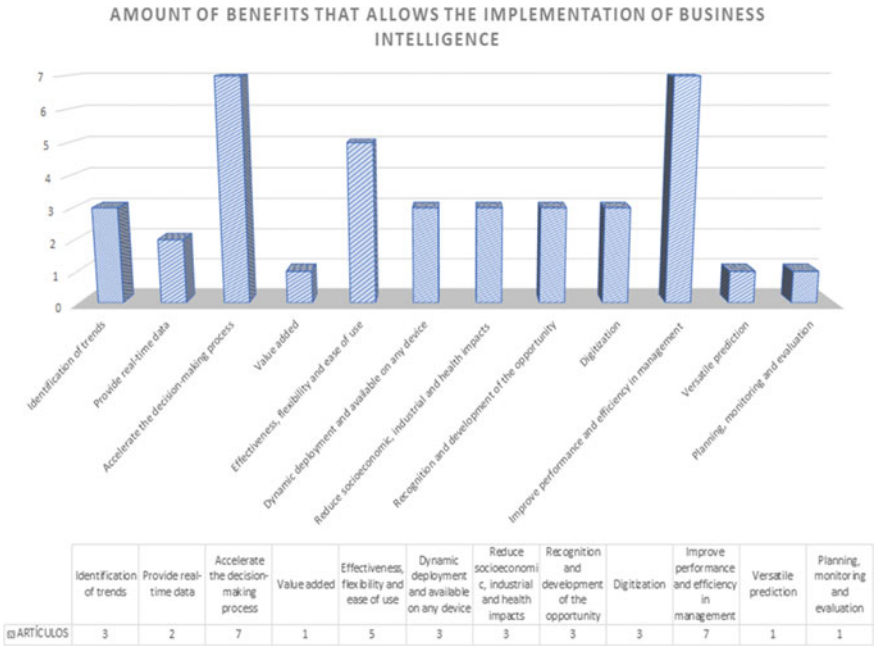


Fig. 6 Number of technologies mostly applied in business intelligence per item



	Identification of trends	Provide real-time data	Accelerate the decision-making process	Value added	Effectiveness, flexibility and ease of use	Dynamic deployment and available on any device	Reduce socioeconomic, industrial and health impacts	Recognition and development of the opportunity	Digitization	Improve performance and efficiency in management	Versatile prediction	Planning, monitoring and evaluation
ARTÍCULOS	3	2	7	1	5	3	3	3	3	7	1	1

Fig. 7 Number of benefits of implementing business intelligence per item

Among them to highlight the stock market prediction, helping analysts to use supporting information about the future of the market, relying on Deep Learning technology to perform stock market prediction of sectors [19].

Table 2 Strategies according to the results obtained

Strategies	Articles
Identification of clusters and patterns	[2–6]
Generate structured data	[2, 5, 7, 8]
Ongoing initiatives to grow the business	[4, 5, 7]
Key performance indicators	[9, 10]
Integrating databases to cloud platforms	[7]
Taking the customer into account and adapting to his preferences	[4, 5, 7, 9, 11, 12]
Technological tools	[13]
Agent-based modeling and simulate effects of measures taken	[14]
Qualitative methods	[4]
Perform analysis of various scenarios	[2, 10, 15–17]

Table 3 Technologies according to the results obtained

Technologies	Articles
Big Data	[5, 16, 18]
Business Intelligence	[5, 12, 19]
Artificial Intelligence	[6, 11, 12, 15, 16, 18, 20, 21]
Machine Learning	[4–6, 11, 21, 22]
Text Mining	[5]
Deep Learning	[19]

Performing an analysis of various scenarios, goes hand in hand with stock market prediction, in order to support strategic decision making. This can provide knowledge through Big Data and Artificial Intelligence technologies, to decision making processes in unpredictable environments, both to searchers and professionals in capturing risks and opportunities of new technologies in knowledge management processes [16].

RQ2. What have been the benefits of Business Intelligence in times of COVID-19 for decision making in the business sector?

According to Figs. 6, 7 and Table 4, it can be seen that the articles related to our topic have benefited from trend identification, accelerate the decision making process, reduce socio-economic, industrial and health impacts, stock market prediction, etc. This result indicates that these benefits resulted in strategic business decision making in COVID-19 times.

The highlight is that it allows you to observe the data in real-time, thanks to the help of different Machine Learning techniques to predict. This provides an easy-to-use and low-cost tool. Also that this could be given on mobile devices, so it would be very useful for decision making [8].

Table 4 Benefits according to the results obtained

Benefits	Articles
Identifying trends	[2, 3, 5]
Providing real-time data	[3, 12]
Accelerate the decision making process	[3, 5, 10, 12, 16, 21, 23]
Added value and innovation	[10, 21]
Effectiveness, flexibility and ease of use	[2, 5, 8, 11, 12]
Dynamic deployment and available on any device	[2, 5, 8]
Reducing socio-economic, industrial and health impacts	[1, 10, 21]
Opportunity recognition and development	[10, 18, 24]
Digitization	[9, 25–27]
Improve management performance and efficiency	[3, 10, 16, 18, 19, 23, 25]
Stock market prediction	[19]
Planning, monitoring and evaluation	[17]

On the other hand, companies are taking advantage of the crisis, seeing it as an opportunity and taking advantage of it, adapting to the needs of the market. Allowing companies opportunities for improvement, saving costs and maintaining their innovation activities [10].

RQ3. Which countries have developed the most research on the topic of Business Intelligence in strategic decision making in times of COVID-19?

According to Fig. 4, it can be seen that there is much more research related to our topic coming from countries such as Germany with the highest amount, followed by India and Italy. This result indicates that there is more research for strategic business decision making in times of COVID-19. The articles related to our topic come from the continents of Europe and Asia (from highest to lowest). This result indicates that there is knowledge in the mentioned continents according to the strategic business decisions in times of COVID-19.

This leads to the fact that most of the research on strategic business decision making takes place in the European continent, followed by Asia. Therefore, much of the knowledge of these strategies and technology support is reflected in their research, which was rescued by the systematic review.

RQ4. Which technologies other than Business Intelligence support strategic business decision making in times of COVID-19?

According to Fig. 5 and Table 3, it can be observed that the articles related to the topic use Big Data, Business Intelligence, Artificial Intelligence, Machine

Learning, Text Mining and Deep Learning technologies. This result indicates that these technologies support strategic business decision making in times of COVID-19.

According to the systematic review of the literature much of the technology support in strategic business decision making is Artificial Intelligence (AI) bringing about changes in society and economy, it provides services and evaluates opportunities. This technology is based on Machine Learning, which is being put into practice. A survey of more than 300 executives was conducted, and the result was that 95% of the participants believe that AI will play an important role in their companies and that it will generate great added value [21].

5 Conclusions

The strategies that have been used for Business Intelligence in times of COVID-19 for decision making in the business sector are identification of clusters and patterns, generating structured data, continuous initiatives to grow the business, key performance indicators, integrating databases to cloud platforms, taking into account the customer and adapting to their preferences, technological tools, agent-based modeling and simulating effects of the measures taken, qualitative methods and performing analysis of various scenarios.

Likewise, the benefits to be highlighted when applying these strategies are trend identification, providing real-time data, accelerating the decision making process, digitalization, improving performance and management efficiency, stock market prediction, planning, monitoring and evaluation.

On the other hand, all our research was based on the years 2020 and 2021 according to our inclusion and exclusion criteria as can be seen in Table 1. Following this, a large part of the research was conducted in the year 2020 in excess of 15 articles.

It can be said that the application of technology in strategic business decision making in times of COVID-19 has been very useful to give continuity to small, medium and large companies, being technology a very important and influential factor in such action. Due to this, positive results are generated that generate benefits within the organizations in the context of the pandemic, because analysis of various scenarios is performed, allowing them to be prepared to face the adversities that may arise, mitigate risks, provide effective solutions for continuous improvement and generate value.

The related work evaluates the techniques used in group decision making to support the risk assessment process. Highlighting that since 2011 they focus mainly on support with information technology, construction and new product development risks. This research, most of their articles come from China, compared to us, they come from Europe, to be more exact from Germany. In different contexts, but always taking into account the risks in order to mitigate them with the help of information technology [28].

For future work, it is recommended to investigate more about the topic specifically applying Artificial Intelligence in strategic business decision making in times of

COVID-19 being a technology that goes hand in hand with Machine Learning, this due to predict possible scenarios that may arise in the future and thus be able to have an idea to anticipate any possible scenario. In addition to this, it is recommended to investigate more about the mentioned strategies since it is there where a greater impact of possible subtopics can be generated that can be taken as a comparison in other studies and thus make a better decision about which would be the best option among them.

References

1. Ghita M, Zineb B, Siham B (2020) Smart pandemic management through a smart, resilient and flexible decision-making system. Int Arch Photogramm Remote Sens Spat Inf Sci ISPRS Arch 44(4/W3):285–294. https://doi.org/10.5194/isprs-archives-XLIV-4-W3-2020-285-2020
2. Sechi GM et al (2020) Business intelligence applied to emergency medical services in the Lombardy region during sars-COV-2 epidemic. Acta Biomed 91(2):39–44. https://doi.org/10.23750/abm.v91i2.9557
3. Aristizábal-Torres D, Peñuela-Meneses CA, Barrera-Rodríguez AM (2020) An interactive web-based dashboard to track COVID-19 in Colombia. Case study: five main cities. Rev Salud Publica 22(2):1–6. https://doi.org/10.15446/rsap.V22n2.87668
4. Fiaidhi J (2020) Envisioning insight-driven learning based on thick data analytics with focus on healthcare. IEEE Access 8:114998–115004. https://doi.org/10.1109/ACCESS.2020.2995763
5. Saritas O, Bakhtin P, Kuzminov I, Khabirova E (2021) Big data augmentated business trend identification: the case of mobile commerce. Scientometrics 126(2):1553–1579. https://doi.org/10.1007/s11192-020-03807-9
6. Devarakonda P, Sadasivuni R, Nobrega RAA, Wu J (2021) Application of spatial multicriteria decision analysis in healthcare: Identifying drivers and triggers of infectious disease outbreaks using ensemble learning. J Multi-Criteria Decis Anal November 2020:1–14. https://doi.org/10.1002/mcda.1732
7. Harris RM (2020) Data warehousing and decision support system effectiveness demonstrated in service recovery during COVID19 health pandemic. In: 2020 14th international conference open source system technology ICOSST 2020—proceedings. https://doi.org/10.1109/ICOSST51357.2020.9333019
8. Awal MA, Masud M, Hossain MS, Bulbul AAM, Mahmud SMH, Bairagi AK (2021) A novel Bayesian optimization-based machine learning framework for COVID-19 detection from inpatient facility data. IEEE Access 9:10263–10281. https://doi.org/10.1109/ACCESS.2021.3050852
9. Vollrath MD, Villegas SG (2021) Avoiding digital marketing analytics myopia: revisiting the customer decision journey as a strategic marketing framework. J Mark Anal 0123456789. https://doi.org/10.1057/s41270-020-00098-0
10. Guderian CC, Bican PM, Riar FJ, Chattopadhyay S (2021) Innovation management in crisis: patent analytics as a response to the COVID-19 pandemic. R D Manag 51(2):223–239. https://doi.org/10.1111/radm.12447
11. Bauer K, Hinz O, van der Aalst W, Weinhardt C (2021) Expl(AI)n it to me—explainable AI and information systems research. Bus Inf Syst Eng 63(2):79–82. https://doi.org/10.1007/s12599-021-00683-2
12. Raineri P, Molinari F (2021) Innovation in data visualisation for public policy making. Springer International Publishing
13. Frp UJ et al (2021) $Qdo\]Lqj &29,' 'Dwdvhw Wkurxjk 'Dwd 0Lqlqj 7Rro 2Udqjh

14. Leitner S, Gula B, Jannach D, Krieg-Holz U, Wall F (2021) Understanding the dynamics emerging from infodemics: a call to action for interdisciplinary research. SN Bus Econ 1(1):1–18. https://doi.org/10.1007/s43546-020-00027-4
15. Liu Y, Lee JM, Lee C (2020) The challenges and opportunities of a global health crisis: the management and business implications of COVID-19 from an Asian perspective. Asian Bus Manag 19(3):277–297. https://doi.org/10.1057/s41291-020-00119-x
16. Iandolo F, Loia F, Fulco I, Nespoli C, Caputo F (2020) Combining big data and artificial intelligence for managing collective knowledge in unpredictable environment—insights from the Chinese case in facing COVID-19. J Knowl Econ. https://doi.org/10.1007/s13132-020-00703-8
17. Al-Dabbagh ZS (2020) The role of decision-maker in crisis management: a qualitative study using grounded theory (COVID-19 pandemic crisis as a model). J Public Aff 20(4):1–11. https://doi.org/10.1002/pa.2186
18. Chaves-maza M, Martel EMF (2020) Entrepreneurship and sustainability issues the policies carried out in the first half of 2020 and proposed by the governments for the recovery of the economy have been very different if we look at the countries of the European Union and associated with it. 8(2):662–681
19. Jabeen A et al (2021) An LSTM based forecasting for major stock sectors using COVID sentiment. Comput Mater Contin 67(1). https://doi.org/10.32604/cmc.2021.014598
20. Thomas O et al (2020) Global crises and the role of BISE. Bus Inf Syst Eng 62(4):385–396. https://doi.org/10.1007/s12599-020-00657-w
21. Buxmann P, Hess T, Thatcher JB (2021) AI-based information systems. Bus Inf Syst Eng 63(1):1–4. https://doi.org/10.1007/s12599-020-00675-8
22. Khakharia A et al (2021) Outbreak prediction of COVID-19 for dense and populated countries using machine learning. Ann Data Sci 8(1):1–19. https://doi.org/10.1007/s40745-020-00314-9
23. Liao H, Tan R, Tang M (2021) An overlap graph model for large-scale group decision making with social trust information considering the multiple roles of experts. Expert Syst 38(3):1–20. https://doi.org/10.1111/exsy.12659
24. Polas MRH, Raju V (2021) Technology and entrepreneurial marketing decisions during COVID-19. Glob J Flex Syst Manag 22(2):95–112. https://doi.org/10.1007/s40171-021-00262-0
25. Guo H, Yang Z, Huang R, Guo A (2020) The digitalization and public crisis responses of small and medium enterprises: implications from a COVID-19 survey. Front Bus Res China 14(1):1–25. https://doi.org/10.1186/s11782-020-00087-1
26. Ding D, Guan C, Chan CML, Liu W (2020) Building stock market resilience through digital transformation: using Google trends to analyze the impact of COVID-19 pandemic. Front Bus Res China 14(1). https://doi.org/10.1186/s11782-020-00089-z
27. Tupia-Astoray A, Andrade-Arenas L (2021) Implementation of an e-commerce system for the automation and improvement of commercial management at a business level. Int J Adv Comput Sci Appl 12(1):672–678. https://doi.org/10.14569/IJACSA.2021.0120177
28. Munhoz Arantes RF, Ribeiro Carpinetti LC (2019) Group decision making techniques for risk assessment: a literature review and research directions. In: IEEE international conference fuzzy system, vol 2019-June, no MCDM, pp 1–6. https://doi.org/10.1109/FUZZ-IEEE.2019.8858927

High-Engagement Chinese Digital Public Diplomacy on Twitter

Aliya Nisa Anindita, Rangga Aditya Elias, Tia Mariatul Kibtiah, Eka Miranda, and Aditya Permana

Abstract Digital public diplomacy has become a new method implied by the state to interact with foreign publics in the Internet era, particularly by utilizing social media. One of the most used social media by the state to carry out the interaction is Twitter. This phenomenon has also made communist states like China utilize Twitter in interacting with foreign publics. Therefore, this paper aims to examine how social media posts intended for public diplomacy can reach a high level of engagement and interact with a more significant number of the public abroad. This paper analyzes posts created by six Chines diplomats on social media Twitter by capturing their engagement, sentiment, as well as the subject. The focus of this paper is to discover that most high-engagement posts contain positive sentiment, that which subject becomes popular to rely on each account's audience, and that some engagement takes the form of negative replies.

Keywords Digital public diplomacy · Social media · China · Sentiment analysis

1 Introduction

Previously public diplomacy was carried out through global mass media. However, the Internet era has enabled a more direct line of communication between the government and the foreign public. The use of Web-based platforms, especially blogs and

A. N. Anindita · R. A. Elias (✉) · T. M. Kibtiah · E. Miranda · A. Permana
Bina Nusantara University, Jakarta, Indonesia
e-mail: raaditya@binus.edu

A. N. Anindita
e-mail: aliya.anindita@binus.ac.id

T. M. Kibtiah
e-mail: tia.kibtia001@binus.ac.id

E. Miranda
e-mail: ekamiranda@binus.ac.id

A. Permana
e-mail: aditya.permana@binus.edu

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_40

social media, for public diplomacy purposes has been called digital public diplomacy or simply digital diplomacy [1, 2].

It has been shown that communication through social media can influence public opinion [3], yet the success of digital public diplomacy has been challenging to measure. One most used method is through assessing a social media post's engagement metrics, which in various social media platforms can be seen in the form of the number of "likes," "reshares," and "views." These metrics represent how far the post has reached, with higher numbers implying that more people have seen the post. Various studies have used these statistics to measure digital diplomacy effectiveness [2, 4–6].

Although measuring engagement metrics is relatively simple, research on how this metric is achieved is still sparse, particularly in the specific context of digital public diplomacy. Therefore, this paper intends to further fill this gap by studying Chinese diplomats' traits shared by successful public diplomacy posts on Twitter. This paper then aims to answer: What traits are shared across tweets by Chinese diplomats on Twitter with the highest engagement?

Past research by Bjola and Jiang [7] proposed that there are three aspects to public diplomacy engagement: (1) agenda setting, referring to the information being disseminated, (2) presence expansion, referring to how many people the information manages to reach, and (3) conversation generating, referring to its use for two-way communication. Their paper's study of posts by E.U., United States, and Japan's embassies in the Chinese social media platform Weibo show that the first two are indeed being done well, though conversation generating was lacking.

Moreover, Khan et al. [4] studied the social media activity of @KoblerInPak, a German ambassador to Pakistan, who was especially popular on his platform of choice: Twitter. They defined engagement as the number of likes, reshares (or "retweets" in the platform's terminology), and replies and proposed that three traits are present in high-engagement tweets: self-disclosure (discussion of personal life and day-to-day events), positive attitude, and inquisitiveness. Analysis of the ambassador's tweets shows that these traits exist in his higher-engagement tweets.

Furthermore, Ittefaq [5] studied India and Pakistan's public diplomacy efforts on the social media platform Facebook and discovered weak public reception. This results from, they argued, the lack of care in creating two-way communication and a confusing mix of domestic and international subjects, resulting in unclear messages and uncertainty of who the intended audience is. Additionally, a study by Yoon and Chung [6] on the European Union's public diplomacy efforts toward South Korea shows low engagement for similar reasons, i.e., the lack of a straightforward narrative in their posts.

Based on these previous studies, this research formulated the hypothesis that social media posts with higher engagement will have the following two traits: (1) discuss subjects that are relevant to their audience and (2) contain positive sentiment. Thus, this paper analyzes social media posts through their subject and sentiment and how these two traits relate to their engagement.

2 Method

Data on this paper was gathered from tweets created by the accounts of six Chinese diplomats during the six months of April 2021 to September 2021. The six-month range was chosen to receive a general scope of a Twitter account's activity, and the period was chosen due to recentness at the time of writing.

This research chooses those accounts in Table 1 because they actively tweet during the time period and have many followers. The focus is on personal instead of embassy accounts because they tend to be more active and receive higher engagement than embassy accounts. The chosen diplomats consist of two regional-scoped diplomats, two ambassadors to specific countries, and two spokespersons to the Foreign Ministry, representing different kinds of Chinese diplomats on Twitter. The account of the Chinese ambassador to the United States, Ambassador Qin Gang tweeting from @AmbQinGang, is not included in this analysis as he has only taken the position and thus has only started tweeting in July 2021.

Tweets were gathered using the official Twitter API v2 on November 1, 2021, using a custom script written in Python. Considering the fast-moving nature of Twitter, by this point in time, the number of engagements will no longer change significantly. It then filtered out "retweets" and replies to another tweet as they are not the focus of this paper.

The examination uses the public engagement metric of these tweets, consisting of likes, retweets, quote retweets, and replies. These numbers are included in the data

Table 1 The accounts in this analysis

No.	Twitter handle	Display name	Account description	Followers
1	Spokesperson CHN	Hua Chunying 华春莹	Foreign Ministry Spokesperson, Director General, Information Department, MFA	997,306
2	zlj517	Lijian Zhao 赵立坚	Spokesman and DDG, Information Department, Foreign Ministry, China	1,038,516
3	China2ASEAN	Ambassador Deng Xijun	Ambassador of the People's Republic of China to ASEAN	51,368
4	Amb. Liu Xiao Ming	刘晓明 Liu Xiaoming	Special Representative of the Chinese Government on Korean Peninsula Affairs, Former Chinese Ambassador to Egypt, the DPRK, and the UK	134,095
5	China_Amb_India	Sun Weidong	Chinese Ambassador to India	90,069
6	AmbNong	Nong Rong	Chinese Ambassador to Pakistan	44,093

returned by the official API. "Engagement" and "engagement number" will refer to these public metrics. Additionally, likes and retweets are considered "passive" engagement, as adding to those numbers only requires a single click from the user, while Replies and quote tweets are considered "active" engagement, as creating replies or quoting tweets requires users to share their thoughts publicly.

Engagement numbers are compared between tweets on the same account, as the number of an account's preexisting audience, as seen from follower count, can influence the number of engagements that a tweet can reach. Text data from each tweet is then processed through VADER, a machine learning-trained sentiment analysis tool specifically attuned to sentiments in social media [8]. VADER gives each tweet a compound polarity score between − 1 and + 1. It shows whether the tweet is leaning more negative or positive. A score equal to or more than 0.05 is considered positive, while equal to or less than − 0.05 is considered negative. Tweets that scored less than 0.05 or more than − 0.05 are considered neutral.

In comparing sentiment with engagement numbers, bubble charts are used to map the sentiment to the y-axis and engagement to the size of the bubble. Meanwhile, the x-axis mapped for the time the tweet was posted. This paper also analyzed tweets with high-engagement numbers for their content and compared them with other accounts. Furthermore, this paper examines whether some topics are shared across accounts or receive incredibly high engagements for specific accounts.

3 Findings and Discussions

3.1 General Overview of Engagement Numbers and Sentiments

From the accounts and time range, the median level of engagement is always lower than the average. Only a few tweets manage to receive a notably higher number of engagements. In general, accounts with higher followers also have higher median engagement (Table 2).

Table 2 General engagement numbers for each account

No.	Twitter handle	Total tweets gathered	Number of followers	Average engagement	Median engagement
1	SpokespersonCHN	1375	997,306	847.5	600.5
2	zlj517	696	1,038,516	1466.6	925.5
3	China2ASEAN	1739	51,368	68.7	41.0
4	AmbLiuXiaoMing	122	134,095	158.76	85.0
5	China_Amb_India	124	90,069	366.5	75.5
6	AmbNong	155	44,093	940.2	620

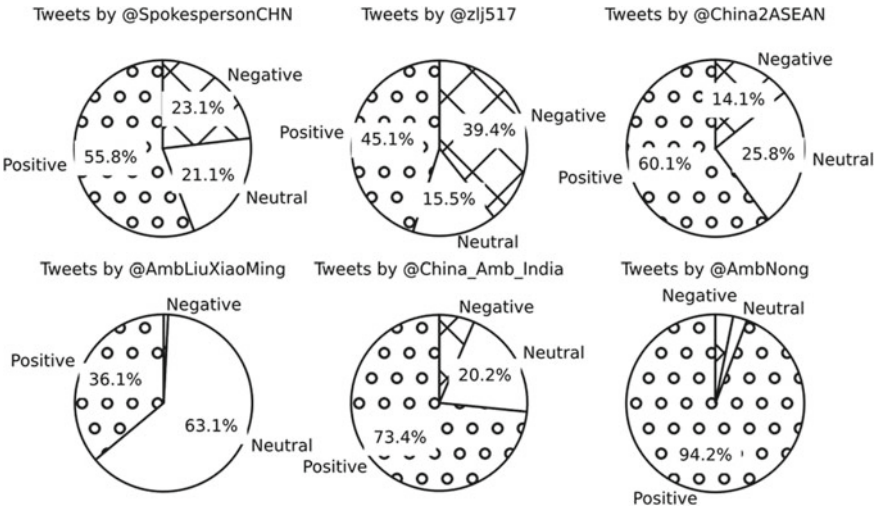


Fig. 1 Pie charts for the sentiments of all tweets in all six accounts

According to the data, most tweets for all accounts are positive, except for @AmbLiuXiaoMing. The account also tweets in Chinese and occasionally posts quote tweets with only an emoji as its text. As VADER Sentiment cannot parse these tweets, they are marked as Neutral. Limiting the sentiment analysis only to English tweets yield 89.8% positive tweets.

Although positive tweets are the majority for every account expected from China's public diplomacy, each account has a different proportion of sentiment. Accounts that tweet for a general, non-regionalized audience, i.e., @SpokespersonCHN and @zlj517, have a lower proportion of positive tweets. Account of ambassador to a specific country, i.e., @China_Amb_India and @AmbNong, creates more positive tweets. The ratio of sentiment for each account is shown in the charts (Fig. 1).

3.2 Accounts Meant for a Global Audience: @SpokespersonCHN and @zlj517

The two accounts in this section are spokespersons to China's Ministry of Foreign Affairs. Their tweets tend to be directed at a more general audience, with discussions of affairs both within China and outside of China.

Figure 2 shows that, just as most of @SpokespersonCHN's tweets are positive, most of their high-engagement tweets are positive. Tweets by @zlj51, as seen in Fig. 3, contain more range in sentiment polarity. However, most high-engagement tweets are still considered positive, especially those with very high engagement.

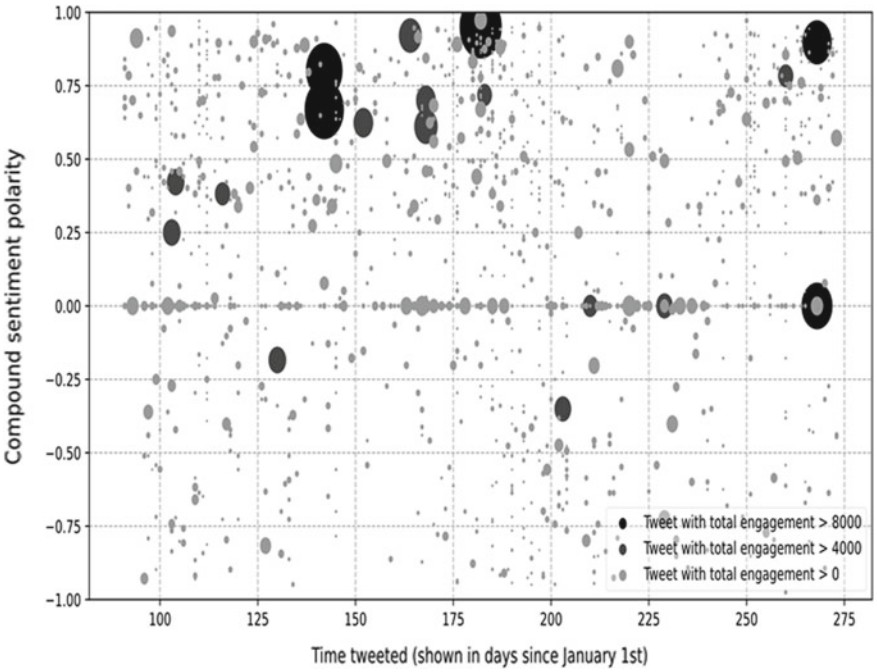


Fig. 2 Bubble chart showing sentiment polarity and engagement number of tweets by @SpokespersonCHN

On both accounts, the highest engagement tweets are celebratory tweets about the 100th anniversary of the China Communist Party, containing very positive sentiment, visible as a large circle in the top-center of the graph. Another topic that receives high engagement in both accounts is those about the death of Yuan Longping, a renowned agronomist. @SpokespersonCHN's tweet is considered very positive, noting the man as a "superstar," while @zlj517's tweets contain more general condolences, a note on how the man "pull countless people out of hunger" and is considered a negative tweet by the sentiment tool.

The one neutral tweet with over 8000 engagements by @SpokespersonCHN contains an image with no accompanying text, showing a woman coming out of an airplane. This tweet celebrates the return to China of Meng Wanzhou, a Chinese businesswoman detained overseas, and thus can be considered a tweet with positive sentiment.

By contrast, @zlj517 tweeted about this subject once, receiving lower engagement than the median. Another contrasting topic is tweeted celebrating China–Pakistan friendship, tweeted by @zlj517 on Pakistan's Independence Day and the anniversary of their diplomatic relation. All these tweets receive high engagement, but the one-time @SpokespersonCHN tweeted about it receives less than median engagement. These two examples indicate that audiences for the two accounts are not very similar after all.

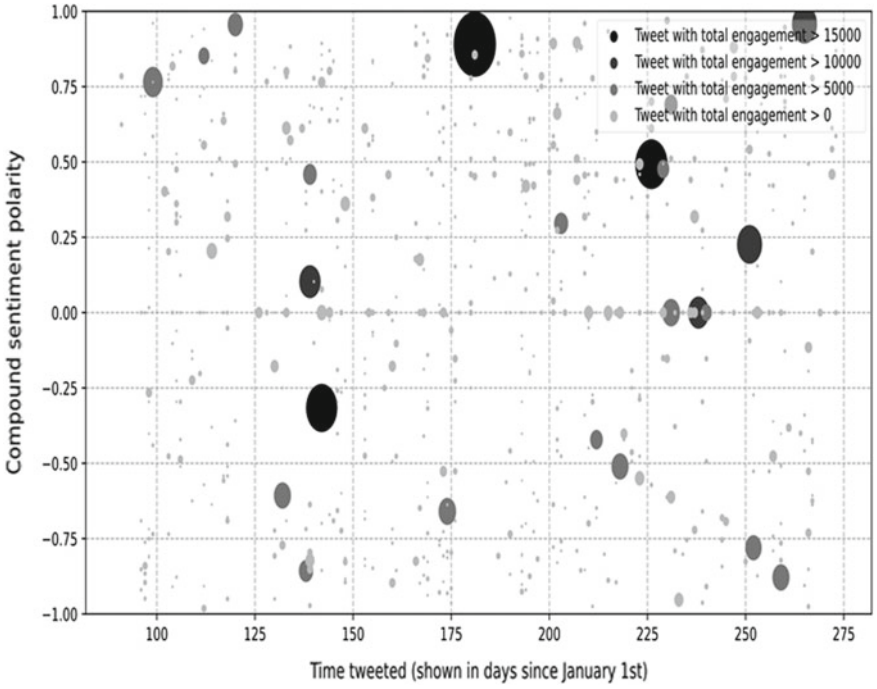


Fig. 3 Bubble chart showing sentiment polarity and engagement number of tweets by @zlj517

3.3 Accounts with Seemingly Regional Scope: @China2ASEAN and @AmbLiuXiaoMing

@China2ASEAN and @AmbLiuXiaoMing are, respectively, the Chinese representative of ASEAN and China's special representative toward Korean affairs. Though they are both similar in their regional scope, data shows that the subjects of @China2ASEAN's tweets are more international, similar to the two accounts from the previous section. Few of his tweets are specifically about or toward Southeast Asians. @AmbLiuXiaoMing, on the contrary, mainly discusses issues relating to the Korean peninsula, both North and South, and personal events (Figs. 4 and 5).

@China2ASEAN is a curious case as, although most of their tweets have positive sentiment, their highest-engagement tweet is negative. The tweet discusses the nuclear wastewater from Fukushima and how it will pollute the Pacific Ocean, accompanied by a short video showing how far the nuclear wastewater will spread.

@zlj517 and @SpokespersonCHN also discussed the subject of nuclear wastewater, but those tweets are not ranked among their highest-engagement tweets, and the higher-engagement ones tend to be positive. As such, this tweet by @China2ASEAN is likely an outlier with factors other than sentiment contributing to its popularity.

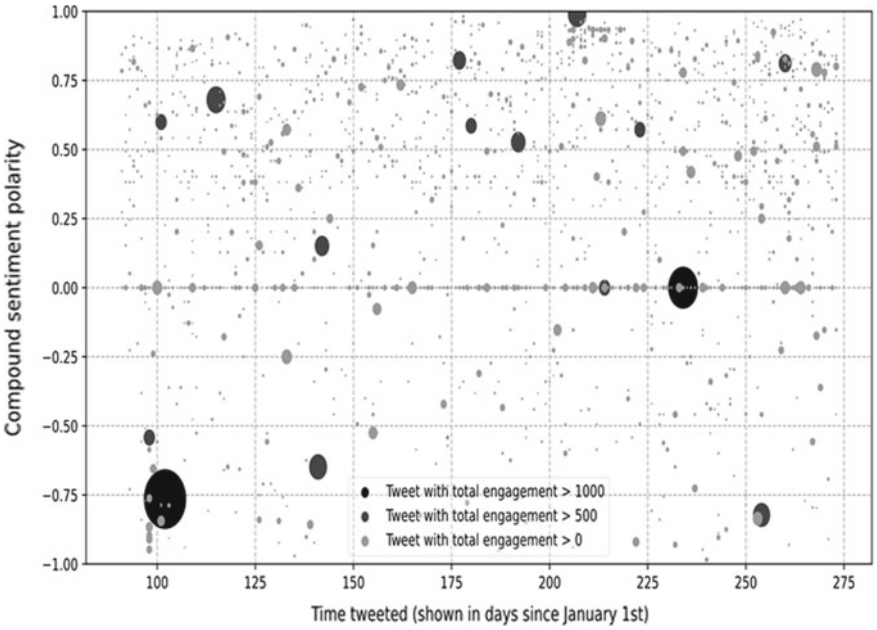


Fig. 4 Bubble chart showing sentiment polarity and engagement number of tweets by @China2ASEAN

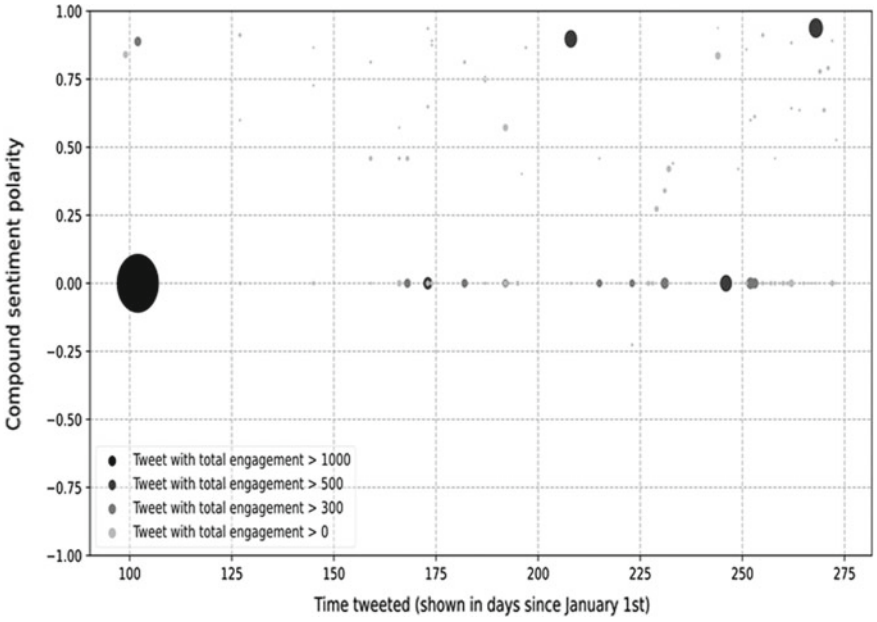


Fig. 5 Bubble chart showing sentiment polarity and engagement number of tweets by @AmbLiuXiaoMing

The topic of wastewater was tweeted four other times by this account and received more than the average, indicating that it is of interest to the account's audience.

The second highest-engagement tweet from this account, which according to the sentiment tool contained neutral sentiment, shows a video that celebrates a Chinese engineering feat, a form of "telling China's story well." A great deal of tweets by @China2ASEAN does the same thing, i.e., promoting China's various accomplishments, and they generally receive higher engagement than the median.

In contrast to @China2ASEAN's international-minded tweets, many @AmbLiuXiaoMing are not in English and thus considered neutral by the sentiment analysis tool. The high-engagement tweets in English are all positive, save for one.

The highest-engagement tweet on this account, which has significantly higher engagement than any other tweet, is written in Chinese and announces the account's holder's new position as China's special representative in the Korean Peninsula affair. Many of the tweet's high-engagement tweets are of this nature, which indicates that a great deal of the account's audience is interested in the account holder's activity, which would align with findings by Khan et al. [4]'s finding that self-disclosure is found in high-engagement tweets.

3.4 Accounts of Specific Country's Ambassadors: @China_Amb_India and @AmbNong

For tweets by @China_Amb_India, there is a vast gap of engagement number between those with high engagements compared to the median level. As shown in the graph in Fig. 6, most of their tweets are barely visible. The visible few receive a vastly higher number of engagements and are clustered in a specific time period. These tweets discuss the transport of Chinese-made medical supplies to India, mainly to combat COVID-19. These tweets contain neutral to positive sentiment, and all receive extensive engagements. It happens as a result of the topic resonating particularly strongly. Of note, a tweet celebrating Indian Independence Day—a positive, celebratory tweet that often receives very high engagement in other accounts—receives relatively lower engagement, though still higher than the average for this account. Among all the accounts in this analysis, tweets by @AmbNong have the best distribution of engagement numbers, i.e., the highest-engagement tweets do not dominate all the engagement numbers, though it still reflects the same trait as all other accounts in that the median is much lower than the average (Fig. 7). Most tweets by @AmbNong are positive, which is also reflected in their high-engagement tweets. The single highest-engagement tweet is considered Neutral by the sentiment tool, but this is only because the tool is not able to parse the customary Muslim greeting for the Eid holidays: "Eid ul-Adha Mubarak!" For reference, the tool gives the equivalent phrase "Happy Easter" the sentiment polarity score of 0.8, which is very positive. Thus, it can be concluded that this account's high-engagement tweets are positive.

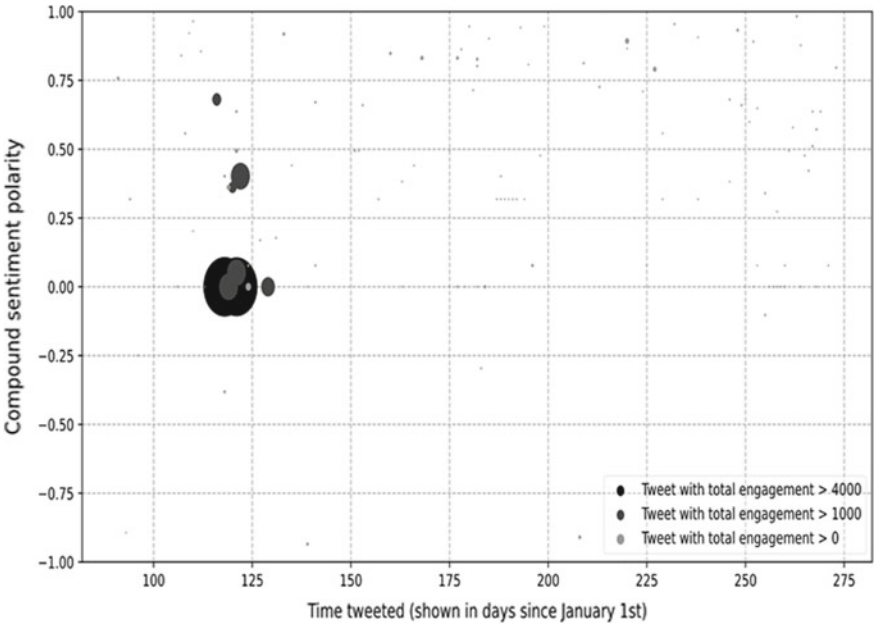


Fig. 6 Bubble chart showing sentiment polarity and engagement number of tweets by @China_Amb_India

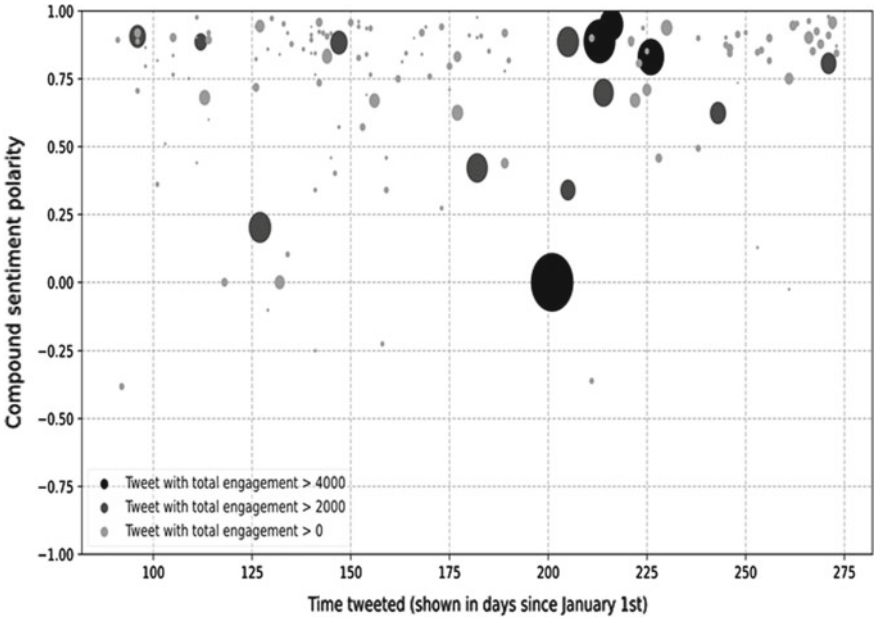


Fig. 7 Bubble chart showing sentiment polarity and engagement number of tweets by @AmbNong

The second highest-engagement tweet congratulated a Pakistani athlete in the Tokyo Olympics. The third highest-engagement tweet celebrates Pakistani Independence Day. In contrast to @China_Amb_India, tweets about pandemic-related aid receive very little engagement. The most commonly tweeted topic in this account is meetings with officials, business people, and journalists, which receive very high engagement, showing that the account's holder and its audience have the same ideas of what kind of tweets are relevant or interesting.

4 Conclusion

Several conclusions can be drawn from the above analysis. First, most tweets by these accounts have positive sentiment, aligning with China's public diplomacy strategy. Second, with few exceptions, tweets with high engagement tend to be positive, even within accounts with a more mixed proportion of negative tweets. Different accounts also tend to have different subjects in their highest-engagement tweets, suggesting different accounts also have different audiences despite their apparent similarity of the target audience. Audience attention also vary between accounts, with some receiving engagements only on specific tweets, while some were having a better general proportion.

Public diplomacy practitioners on social media should pay attention to the kind of audience that comes to their accounts—tailoring their tweets to those that the audience likes will likely yield higher engagement and, therefore, higher spread across more people. However, diplomats should also note that not all engagement numbers are equal. Tweets with a high ratio of active engagement can indicate a high number of negative replies, which can happen even when the tweet contains positive sentiment. These negative interactions can still have positive impact on a tweet's reach because Twitter's algorithm favours any kind of engagement [9], but it can negatively impact an account's reputation.

As this paper is still limited in its scope and methodology, it discovered that some high-engagement tweets cluster together but still could not identify events that might cause them. Future research can take a closer look at those specific matters. Another finding is that some tweets receive an unusually high number of active engagements; future research can look at the replies and the accounts that created them.

References

1. Collins SD, DeWitt JR, LeFebvre RK (2019) Hashtag diplomacy: twitter as a tool for engaging in public diplomacy and promoting U.S. foreign policy. Place Brand Public Diplomacy 15:78–96. https://doi.org/10.1057/s41254-019-00119-5
2. Adesina OS (2017) Foreign policy in an era of digital diplomacy. Cogent Soc Sci 3:1297175. https://doi.org/10.1080/23311886.2017.1297175

3. Weeks BE, Ardèvol-Abreu A, Gil de Zúñiga H (2015) Online influence? Social media use, opinion leadership, and political persuasion. Int J Public Opin Res edv050. https://doi.org/10.1093/ijpor/edv050
4. Khan ML, Ittefaq M, Pantoja YIM, Raziq MM, Malik A (2021) Public engagement model to analyze digital diplomacy on twitter: a social media analytics framework. Int J Commun 15:29
5. Ittefaq M (2019) Digital diplomacy via social networks: a cross-national analysis of governmental usage of facebook and twitter for digital engagement. J Contemp East Asia 18:49–69 (2019). https://doi.org/10.17477/JCEA.2019.18.1.049
6. Yoon S-W, Chung SW (2020) The E.U.'s public diplomacy in Asia and the world through social media: sentiment and semantic network analyses of official facebook pages of European External Action Service and E.U. Delegation to the Republic of Korea. J Contemp East Asia 19:234–263 (2020). https://doi.org/10.17477/JCEA.2020.19.2.234
7. Bjola C, Jiang L (2015) Social media and public diplomacy: a comparative analysis of the digital diplomatic strategies of the E.U., U.S., and Japan in China. In: Digital diplomacy. Routledge, pp 85–102
8. Hutto CJ, Gilbert EE (2014) VADER: A parsimonious rule-based model for sentiment analysis of social media text. Eighth international conference on weblogs and social media (ICWSM-14). Ann Arbor, MI, June 2014
9. Twitter: about your home timeline on twitter. https://help.twitter.com/en/using-twitter/twitter-timeline

EfficientNeXt: EfficientNet for Embedded Systems

Abhishek Deokar and Mohamed El-Sharkawy

Abstract Convolutional neural networks have come a long way since AlexNet. Each year the limits of the state-of-the-art are being pushed to new levels. EfficientNet pushed the performance metrics to a new high and EfficientNetV2 even more so. Even so, architectures for mobile applications can benefit from improved accuracy and reduced model footprint. The classic Inverted Residual block has been the foundation upon which most mobile networks seek to improve. EfficientNet architecture is built using the same Inverted Residual block. In this paper we experiment with Harmonious Bottlenecks in place of the Inverted Residuals to observe a reduction in the number of parameters and improvement in accuracy.

Keywords Convolutional neural network · Coordinate attention · Adaptive sharpness aware minimization · Harmonious bottleneck · Sandglass bottleneck · Inverted residual · Fused MBConv · CIFAR-10

1 Introduction

Convolutional neural networks (CNN), at this point are pretty much ubiquitous in the domain of machine learning. With the availability of cheap computing power and advancements in processing hardware, the improvements to neural network architectures keep coming.

With respect to machine learning, there is an important aspect to focus on, machine learning at edge, which is essentially the idea of deploying ML models on low compute power hardware. The models deployed on such hardware need to be small yet not lose too much of their ability. One of the most import developments in this domain was the introduction of MobileNet [1]. It started off a development into

A. Deokar (✉) · M. El-Sharkawy
Department of Electrical and Computer Engineering, Purdue School of Engineering and Technology Indianapolis, Indianapolis, IN, USA
e-mail: ardeokar@purdue.edu

M. El-Sharkawy
e-mail: melshark@purdue.edu

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_41

reducing the model size while still maintaining respectable levels of accuracy on the commonly used datasets.

Multiple successors were proposed since and MobileNet itself has been worked upon a fair number of times with each successor version achieving significantly better performance. Notably, MobileNetV2 [2] introduced the central idea of Inverted Residual (MBConv) blocks which have become a commonly applied layer in a lot of newer research.

In recent years, the inverted residual bottleneck has been used as the building block for EfficientNet [3] architecture, which made massive strides in accuracy and model size and quickly became the benchmark network for datasets like ImageNet [4], CIFAR10 [5], etc. EfficientNetV2 [6] builds upon EfficientNetV1 by adding a new layer called Fused MBConv, proposed in [7]. The authors found that Depthwise Convolution's performance is slow in the early layers but more effective later [6].

We investigate the performance of Sandglass block introduced with MobileNeXt [8] architecture and Harmonious Bottlenecks introduced in HBONet [9]. This work presents the combinations of layers from MobileNeXt, EfficientNetV2 and HBONet with modifications and proposes a new architecture that utilizes the benefits of the three networks discussed above.

2 Background

In recent years, compute power has increased dramatically. Despite that, prudent research in reducing the compute utilization has driven the ML community to investigate improving neural network performance by decreasing training time and increasing accuracy levels.

EfficientNets are a new family of architectures designed using Network Architecture Search with scaling in mind. The initial approach was formulaic, in that the depth, width and resolution factors were scaled using a formula. EfficientNetV2 improved upon the bottlenecks observed in the development of EfficientNetV1; training on large image sizes was slower, depthwise convolutions in early layers are slow, scaling depth, width and resolution equally was found to be sub-optimal. EfficientNetV2 was then searched for using neural architecture search in a new search space which contained Fused MBConv layer and the resultant V2 was faster to train.

Harmonious bottleneck was proposed in the HBONet paper, they aim to mitigate the increasing consumption of computational resources arising from channel expansion operations by compressing the channel expansion–contraction component into a pair of inverse spatial transformations (the spatial contraction–expansion component) [9] (Fig. 1).

Sandglass blocks were proposed in [8]. Like Harmonious bottleneck and Fused MBConv layer, the Sandglass block was also designed to improve upon the inverted residual block. The design of the block was motivated by the idea of preserving more information between the bottom layers and the top layers. More information is preserved is by adding shortcut connections between the output and input tensors. The

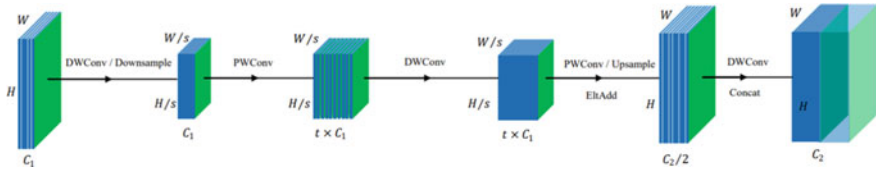


Fig. 1 Harmonious bottleneck operation [9]

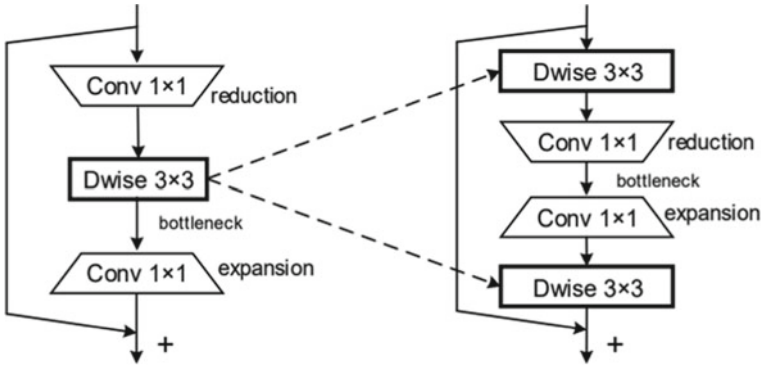


Fig. 2 Sandglass structure [8]

position of depthwise convolutions is also switched such that depthwise convolutions are the outer layers of the sandwich and the inner layers are formed by regular convolution that first reduces and then expands the number of channels [8] (Fig. 2).

3 EfficientNeXt Architecture

In this paper we propose modifications to the architectures mentioned in the background section. In this section we discuss the modifications in detail.

3.1 Layer: Sandglass

Sandglass is effectively a combination of residual bottleneck and inverted residual bottleneck, designed to incorporate the best features of both. It is constructed such that the first and last layers are channel preserving spatial depthwise convolutional layers and sandwiched in between are two pointwise operators that at first reduce the number of channels and then increase them after. This offers a decrease in number of parameters while maintaining performance comparable to MBConv. Coordinate

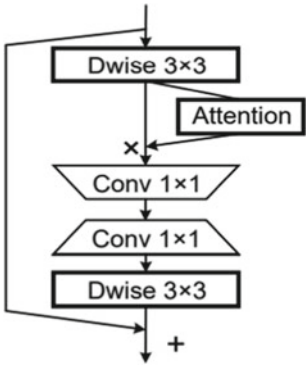


Fig. 3 Sandglass layer with coordinate attention [10]

Attention mechanism is added to the sandglass layer toward the end of the layers as discussed in [10], to improve performance of the Sandglass layer at a slight cost of added parameters (Fig. 3).

3.2 Layer: Fused MBConv

The Fused MBConv layer was designed to utilize the performance offered by a TPU, since depthwise separable convolutions perform slower on TPUs. This performance improvement is achieved by replacing the depthwise convolution and expansion convolution in MBConv with a single regular convolution, reducing the number of activations as compared to MBConv. We use GELU activation in the Fused MBConv block instead of Swish as used in [6].

3.3 Layer: HBNeck

In this paper we take inspiration from [11] and reduce the kernel size in the first depthwise convolution to 3×3 instead of 5×5 as in [9]. This yields reduced number of parameters, in turn reducing model size which is one of the primary motivations of this paper (Fig. 4).

3.4 Activation: GELU

Gaussian Error Linear Unit (GELU) is an activation function, first proposed in [12] and is given by the equation:

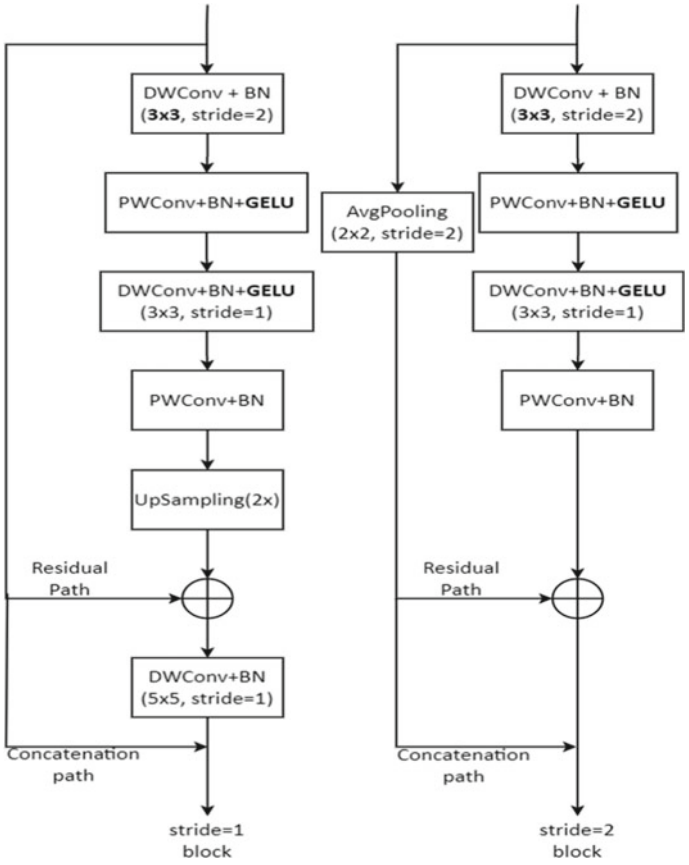


Fig. 4 Modified HBNeck

$$\text{GELU}(x) = x\Phi(x) \tag{1}$$

GELU(x) = x Φ (x)(1)

where $\Phi(x)$ is the standard Gaussian cumulative distribution function (Fig. 5).

The activation functions used in the HBONet, HBONext and EfficientNetV2 architecture are ReLU6 [1], ELU [13] and Swish [14], respectively. We use ReLU6 in the Sandglass layer and experiment with keeping all activations the same in Fused MBConv and Harmonious Bottleneck layers and use GELU, Swish, Mish [15], ReLU6 and ELU and find that GELU activation outperforms the other activations by a considerable margin (Table 1).

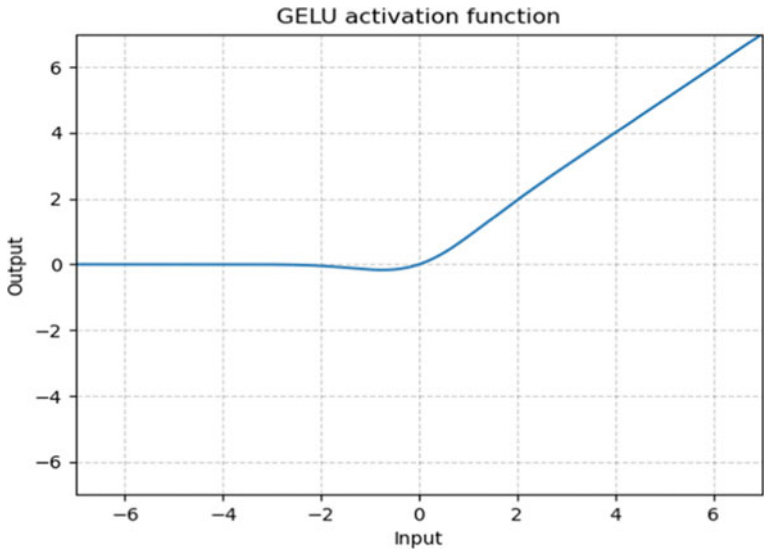


Fig. 5 GELU activation [17]

Table 1 EfficientNeXt architecture

Input size	Layer	*t*	*c*	*n*	*s*
$32^2 \times 3$	Conv2D, k3 × 3	1	64	1	2
$16^2 \times 64$	Sandglass	2	32	1	1
$16^2 \times 32$	Fused MBConv	1	16	1	1
$16^2 \times 16$	Fused MBConv	4	32	2	2
$8^2 \times 32$	Fused MBConv	4	64	4	2
$4^2 \times 64$	HBNeck	4	96	4	1
$4^2 \times 96$	HBNeck	2	128	4	2
$2^2 \times 128$	HBNeck	2	256	4	1
$2^2 \times 256$	Conv2D, k1 × 1	1	512	1	2
$1^2 \times 512$	Sandglass, CA	2	256	2	2
$1^2 \times 256$	Sandglass, CA	2	128	1	1
$1^2 \times 128$	Sandglass	1	10	1	1
$1^2 \times 16$	Conv2D, k1 × 1	–	1600	1	1
$1^2 \times 1600$	AdaptiveAvgPool	–	–	1	–
1 × 10	FC Layer	–	K	–	–

Table 2 Performance comparison with baseline

Model	Parameters (M)	Model size (MB)	Accuracy (%)
EfficientNeXt (width = 1.5)	5.6	22.8	**92.92**
EfficientNet (width = 1.5)	13	52	91.97
EfficientNeXt (width = 1)	2.5	10.5	**92.13**
EfficientNet (width = 1)	5.9	24	90.76
EfficientNeXt (width = 0.75)	1.4	**6.2**	**90.99**
EfficientNet (width = 0.75)	3.5	14.5	89.76

4 Training Setup

- Apollo 6500 with an Intel Xeon Gold 6248 CPU
- Nvidia Tesla V100-PCIE-32 GB
- CUDA 11.2
- Python 3.9.6
- Pytorch 1.9.0.

We train the network on CIFAR10 dataset with a standard train-test split. We used the SGD optimizer with a weight decay of 5e−4, a momentum of 0.9, and a learning rate of 0.02. We use Adaptive Sharpness Aware Minimization (ASAM) [16] which adds 1% improvement in accuracy with some additional training time with rho of 0.5 and eta of 0. The learning rate scheduling used is Cosine Annealing. We use a batch size of 128 and train the network for 200 epochs. Augmentations added to the dataset include random horizontal flipping, cropping and normalizing with CIFAR10 mean.

For fairness, baseline was also trained on the improved training step and the metrics obtained and listed in Table 2 reflect those training improvements.

5 Results

We train our proposed model EfficientNeXt from scratch on CIFAR10 dataset, observe improvements on the accuracy of EfficientNetV2-B0 model with final accuracy of 92.13% while reducing the model size by 56.25%. EfficientNet are networks primarily trained on ImageNet dataset and the number of classes in ImageNet is 1000. In the network definition section of the model, specifying the number of classes as 1000 for EfficientNet yields number of parameters to about 7.1 M parameters as in

[6] and model size of 29.1 MB; the accuracy for EfficientNet with 7.1 M parameters on CIFAR10 dataset is about 90.93% with our training method (Figs. 6 and 7).

Our proposed model scaled down to 0.75 width, slightly outperforms baseline with a model size of 6.2 MB and 1.4 M parameters, a reduction of 74.16 and 76.27% when compared to baseline model size and number of parameters, respectively.

We also compare the performance of the network with other activation functions namely, Mish, Swish, ReLU and ELU (Fig. 8).

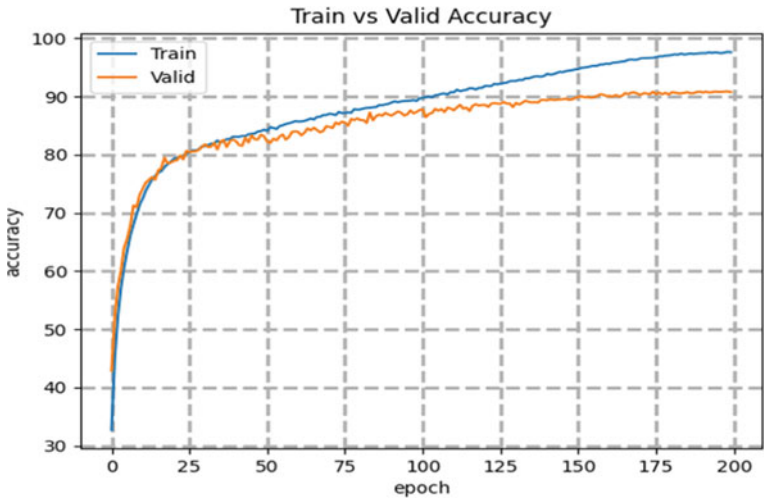


Fig. 6 EfficientNetV2-B0 performance

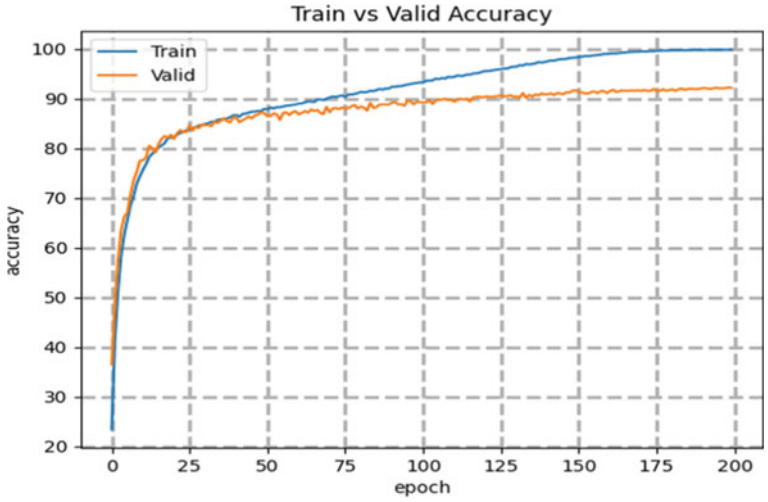


Fig. 7 EfficientNeXt performance

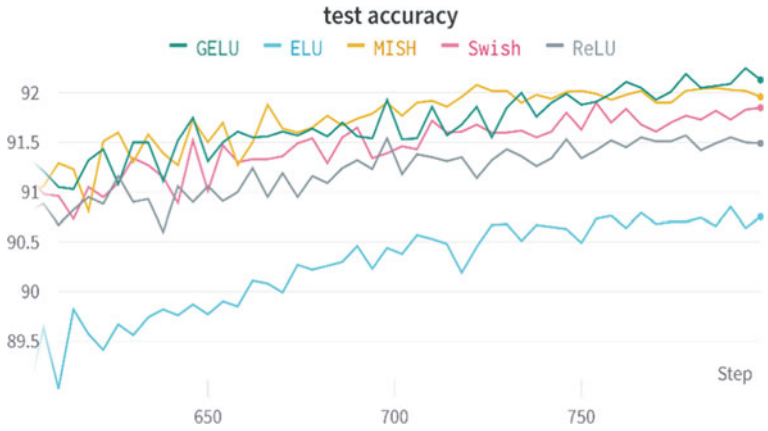


Fig. 8 Activation function comparison [18]

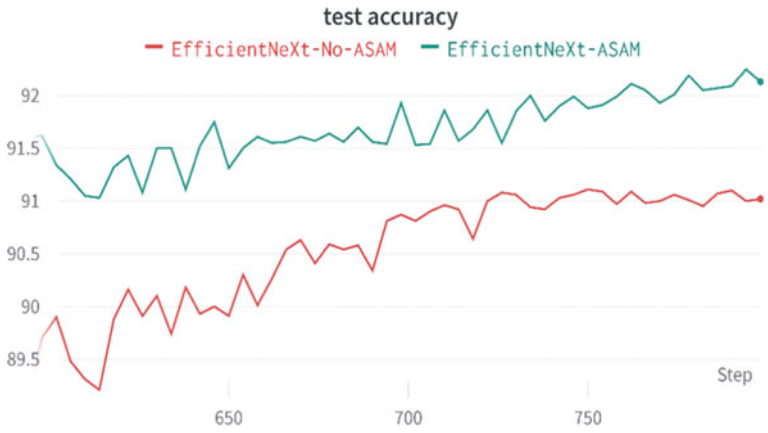


Fig. 9 ASAM performance [18]

We observe a significant improvement with ASAM training step and the results of the difference in accuracy using regular training and ASAM training are shown in Fig. 9. We observe a gain of slightly more than 1% point with ASAM.

6 Conclusion

We introduced EfficientNeXt in this paper and observe a marked improvement over the baseline EfficientNetV2-B0 in terms of accuracy, model size and a sizeable reduction in the number of parameters. Our adoption of newer training strategy based on sharpness aware minimization led to significant gains in accuracy.

With respect to mobile deployment of the model described here, future work could benefit from methods of pruning and quantization to possibly reduce model size further. We do not investigate hyperparameter tuning in this paper, future work could also benefit from the same.

Transfer learning is another approach that holds potential to improve accuracy of the EfficientNeXt model along with Auto Augmentation [19] of the dataset.

Acknowledgements This research was supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute [20].

References

1. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Adam H (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861
2. 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
3. Tan M, Le Q (2019, May) Efficientnet: rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR, pp 6105–6114
4. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Fei-Fei L (2015) Imagenet large scale visual recognition challenge. Int J Comput Vis 115(3):211–252
5. Krizhevsky A, Nair V, Hinton G (2009) Cifar-10 and cifar-100 datasets. 6(1):1. https://www.cs.toronto.edu/kriz/cifar.html
6. Tan M, Le QV (2021) Efficientnetv2: smaller models and faster training. arXiv preprint arXiv:2104.00298
7. Gupta S, Tan M (2019) Efficientnet-edgetpu: creating accelerator-optimized neural networks with automl. https://ai.googleblog.com/2019/08/efficientnet-edgetpu-creating.html
8. Zhou D, Hou Q, Chen Y, Feng J, Yan S (2020) Rethinking bottleneck structure for efficient mobile network design. In: Computer vision–ECCV 2020: 16th European conference, Glasgow, UK, 23–28 August 2020, Proceedings, Part III 16. Springer International Publishing, pp 680–697
9. Li D, Zhou A, Yao A (2019) Hbonet: harmonious bottleneck on two orthogonal dimensions. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 3316–3325
10. Hou Q, Zhou D, Feng J (2021) Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 13713–13722
11. Joshi SR, El-Sharkawy M (2021, June) HBONext: HBONet with flipped inverted residual. In: 2021 IEEE international conference on design and test of integrated micro and nano-systems (DTS). IEEE, pp 1–5
12. Hendrycks D, Gimpel K (2016) Gaussian error linear units (gelus). arXiv preprint arXiv:1606.08415
13. Cheng Q, Li H, Wu Q, Ma L, Ngan KN (2020) Parametric deformable exponential linear units for deep neural networks. Neural Netw 125:281–289
14. Ramachandran P, Zoph B, Le QV (2017) Searching for activation functions. arXiv preprint arXiv:1710.05941
15. Misra D (2019) Mish: a self regularized non-monotonic activation function. arXiv preprint arXiv:1908.08681

16. Kwon J, Kim J, Park H, Choi IK (2021) ASAM: adaptive sharpness-aware minimization for scale-invariant learning of deep neural networks. arXiv preprint arXiv:2102.11600
17. Pytorch: GELU. https://pytorch.org/docs/stable/generated/torch.nn.GELU.html
18. Biewald L Experiment tracking with weights and biases. Weights & Biases. [Online]. Available http://wandb.com/. Software available from wandb.com
19. Cubuk ED, Zoph B, Mane D, Vasudevan V, Le QV (2018) Autoaugment: learning augmentation policies from data. arXiv preprint arXiv:1805.09501
20. Stewart CA, Welch V, Plale B, Fox G, Pierce M, Sterling T (2017) Indiana University Pervasive Technology Institute, Bloomington, IN.https://doi.org/10.5967/K8G44NGB

Discretization and Representation of a Complex Environment for On-Policy Reinforcement Learning for Obstacle Avoidance for Simulated Autonomous Mobile Agents

Andreas Dyrøy Jansson

Abstract In recent years, the demand for digitalization, automation, and smart systems in the airline industry has accelerated. Furthermore, due to the ongoing global pandemic as of 2022, airlines are faced with the challenge of offering flexibility in both cargo and passenger capacity. Studies show that the use of smart products and autonomous agents are expected to play a key part in the digital transformation of the logistics industry. This paper aims to examine the current state-of-the-art in multi-agent systems and reinforcement learning with special interest in intelligent baggage handling systems. How to simplify, implement, and simulate a system of autonomous baggage carts as a software model in order to examine congestion situations will be the main topics of this paper. Furthermore, how the findings from the software model may be applied to real-world scenarios related to Industry 4.0, and baggage handling will also be discussed.

Keywords Autonomous agents · Congestion · Discretization · On-policy · Reinforcement learning · Self-optimization · Simulation

1 Introduction

Due to the outbreak of Covid-19, aviation demand has fallen considerably. According to [1], leisure trips and tourism were first to recover after previous global crises and are expected to surpass business travel demand as the pandemic subsides. During this crisis, airlines have been taking major financial hits, amassing more than $180 billion worth of debt in 2020 [1]. As a result, the industry as a whole is required to cut costs by restructuring for greater efficiency. This means investing in IT, digitalization, and automation. As of 2019, investing in business intelligence (BI) and smart systems was a top priority for 95% of airlines and 87% of airports [2].

A. D. Jansson (✉)
Department of Computer Science and Computational Engineering, UiT The Arctic University of Norway, 8514 Narvik, Norway
e-mail: andreas.d.jansson@uit.no

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_42

Furthermore, due to increased globalization, competition and e-commerce, international competitiveness of companies will depend on effective and on-time delivery of products globally. Similarly, global supply chains are dependent on air freight to reach their customers as quickly and efficient as possible. However, as many passenger flights were grounded during the pandemic, cargo capacity was also reduced [1] estimated that passenger flights are responsible for half of total air cargo capacity, which means that air freight will be a limited resource going forward. As a consequence, airlines need to be agile and grow cargo capacity while still being flexible and able to quickly adapt to changing demand. As a result, previous efforts in digitization have been accelerated and are expected to play a vital part in this transformation [1, 2].

Sensors enable smart products to detect and report problems and, in some cases, even make decisions on their own [3]. Smart products appear in several forms and in various domains, and one of the best known is perhaps the automated guided vehicle (AGV). According to Sella et al. [4], using AGVs for resource transportation may help improve effectiveness, safety, and flexibility in the context of Industry 4.0, and by extension, baggage handling and logistics. To achieve this, agents representing such AGVs are required to have a solid strategy for avoiding obstacles in a complex and dynamic environment. However, a limiting factor that must be considered is the computing performance of the AGVs onboard computers. As such, the main topic this paper will examine is how a complex, dynamic environment can be broken down into discrete states to run on low-power hardware while still being useful for obstacle avoidance. This will be done using a simulation of multiple autonomous agents learning an obstacle avoidance strategy through on-policy reinforcement learning. A simple prototype application to simulate and visualize the agents will be developed and tested, and the results are discussed. How agents are able to avoid dynamic obstacles, in addition to changes in available space will be examined and discussed. Finally, how the findings from this paper may be adapted in order to be useful for real-world challenges regarding baggage handling, and logistics will be brought up.

2 Similar Work and Experiments

The use of multi-agent systems (MASs) and reinforcement learning (RL) in the industry has been documented thoroughly. Examples include [5–7], where smart systems, Internet of Things (IoT), and auction-based task allocation was used to decentralize control of jobs on the factory shop floor. As the manufacturing industry moves toward a more customer-driven market, companies are required to shorten product life cycles and reduce time to market without negatively impacting quality and costs. Such a shift demands more decentralized, flexible control, and increased robustness. Findings showed that these systems were robust in the face of disruptions and unpredictable events, due to their decentralized control.

Experiments with reinforcement learning in combination with MAS have been performed on the logistical challenge of air traffic flow management in [8]. The authors pointed out that the current air traffic control was slow to respond to changing weather conditions and other unexpected events due to centralized control. They proposed a system of agents who were assigned specific 2D points throughout the airspace, in which each agent was tasked with keeping the separation of incoming traffic to a required distance.

Finally, various approaches for route planning for autonomous guided vehicles has been examined. Examples include [9], which focused on creating a fast algorithm able to run on hardware with limited computational power. This was achieved by creating an encoded abstraction of the agent's environment, reducing the size of the input vector. Sensor data and previously generated routes were collected and used to build the model to find the optimal route. This model had to be maintained over time, and a long data collection period was required in order to aggregate sufficient training data.

Similarly, [10] showed how LiDAR sensor data combined with odometer data could be used to determine a robot's position in an unknown environment. However, the major drawback of this approach was that a manually constructed map created up front was required.

A similar approach examined a LiDAR-only based navigation algorithm for a weeding robot navigating in an unknown environment [11]. Information about the environment was collected and used to construct a model in real time, which the robot utilized in order to move along detected lines in rows of crops. Their approach was promising given the limited data available to the robot, and where no prior information about an environment was available.

The use of visual information combined with LiDAR scans to improve accuracy has been discussed in [12]. Using an extensive set of pre-collected images, a robust and accurate strategy for the robot was created. Extensive manual effort required up front was a major drawback, which also meant that changes in the environment could throw the robot off.

In summary, we see many examples of MAS and RL in the industry, both for baggage handling and logistics, and obstacle avoidance and navigation for autonomous agents. However, for the work done with regards to robotics and autonomous agents, the main focus has been on single agents, in environments with few obstacles [7] did address a system with 300 agents, but they were not dynamic, rather they served the role of mediators controlling resource flow throughout the BHS. Furthermore, the state space in works like [9–12] was relative to the static environment, and not to the agent itself, making for a less flexible system. On the other hand, [8] used a system with rewards and state space relative to the agents, which was able to adapt quickly to unexpected events. The agents in [9] used a centralized approach, where the model had to be monitored and re-trained regularly if performance began to decline.

Based on these findings, this paper aims to examine how a highly complex, multi-agent system can be broken down and represented using a simple, tabular approach, with no human interaction. Furthermore, control will be bottom-up and decentralized, and with no information about the environment required up front.

3 Approach and Implementation

It was decided to examine a model of multiple autonomous, mobile agents in an environment consisting of a large number of obstacles (200+). A simple obstacle avoidance simulator based on off-policy RL developed in C# and Windows Forms in a previous work [13] was extended, and new benchmarks were defined.

3.1 State Space and Action Definition

The simulation should consist of a simplified model, while still having some relevancy to real-world applications. This was achieved by using a discretized state space, where the current state of an agent was defined by its distance and angle relative to its nearest obstacle, somewhat similar to [8]. The state space was constructed using two different discretization approaches:

Relative Angle and Distance to Nearest Obstacle. In the first approach, the state was based on the distance and relative angle to the nearest obstacle in range. Distance and angle values were discretized to construct the state space, defined as a $2D$ matrix of width W and height H. W and H thus determined the resolution of the state space in each dimension, which will be discussed later. Sensor values were discretized to a' and d' using the following formulae:

$$a' = \frac{a}{360} \times (H - 1) \tag{1}$$

$$d' = \frac{d}{100} \times (W - 1) \tag{2}$$

This is visualized in the figure below:

In (1), the measured angle value a is discretized into intervals of 15°. Similarly, in (2), the measured distance d is divided by the total sensor range of 100 pixels and discretized into intervals of 5 pixels. Figure 1 (a) shows a visual representation of the state space, as seen by the sensor of a square agent, and the corresponding discrete state space (b). In the current implementation, H was set to 25, and W was set to 20. This produced a $2D$ state space matrix of size $25 \times 20 = 500$ possible states.

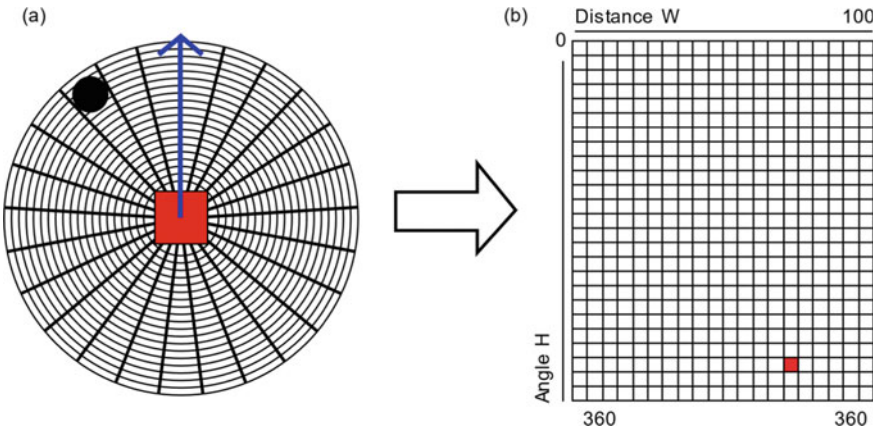


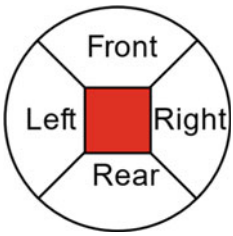
Fig. 1 State space discretization example 1. Current agent orientation and heading is shown by the blue arrow

Cardinal Partitions and Number of Obstacles in each Partition. In the second approach, the state of the agent was defined based on relative angle and the number of obstacles in each interval. Instead of discretizing the angles into 25 15-degree intervals, four 45-degree partitions based on the cardinal directions were manually defined:

A visual representation of the state space discretization and partitioning is shown in Fig. 2. Next, the actual state was determined based on the presence of one or more obstacles in each quadrant. The ideal state would thus be no obstacles in the front partition, where the agent is heading. Similar to the first approach, states were mapped to indices in a $1D$ state matrix by the following method:

Figure 3 shows how indices were generated based on the presence of obstacles in each of the quadrants. This is simply a true or false value, represented as 0 and 1, going clockwise from the front partition. Since there were four partitions of two "sub-states" each, the total size of this state matrix became $2^4 = 16$. Indices of the state matrix were then generated by converting from the binary to decimal value.

Fig. 2 State space partitions based on cardinal directions



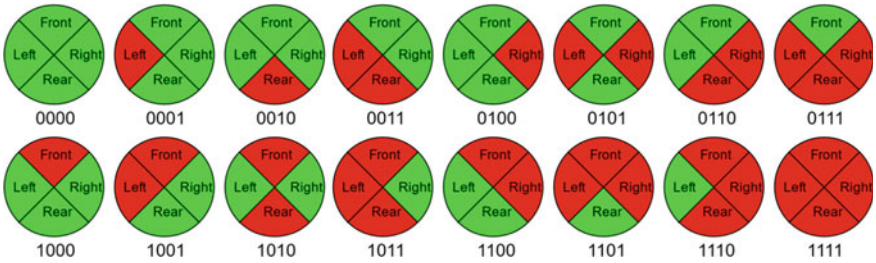


Fig. 3 Partition state mapping visualization

Action Definition. In every state of both approaches, an agent could take one of four actions A': turn left, turn right, wait, or keep going. The actions for turning left and right were performed in the simulator by rotating the agent's direction vector $\vec{V}$ by a fixed amount of 10° in either direction. The wait action was implemented by not updating the agent's position for 10 simulator ticks. As the simulator ran at 125 ticks per second, the agent would wait for 1250 ms.

3.2 Reinforcement Learning Implementation

The reinforcement learning algorithm used was SARSA, which is an on-policy reinforcement learning algorithm [14].

Next-State Estimation. In order to determine the next state following the selected action, "look-ahead" functions for each of the two approaches were implemented. For the first approach, its current trajectory according to the selected action A was used, as shown in Fig. 4. The result from this computation was in turn used to estimate the next distance and relative angle $d_t + 1$ and $a_t + 1$, and thus $S_t + 1|A$:

The agent is shown in Fig. 4 as a green square, with an obstacle shown as a black circle. The current distance d_t and relative angle a_t were provided by the sensor. $\left|\vec{V}\right|$ was also known, as this was the constant velocity of all agents. The next estimated position given action A was then found by:

$$(x, y)_{t+1} = (x, y)_t + \overrightarrow{V}|A, \quad \vec{D}_{t+1} = \left[x_O - x_{t+1}, y_O - y_{t+1}\right] \tag{3}$$

Based on (3), d_{t+1} and a_{t+1} could be calculated:

$$d_{t+1} = \sqrt{x_D^2 + y_D^2}, \quad a_{t+1} = \cos^{-1} \frac{\vec{V} \cdot \overrightarrow{D_t + 1}}{\left|\vec{V}\right| \times \left|\overrightarrow{D_{t+1}}\right|} \tag{4}$$

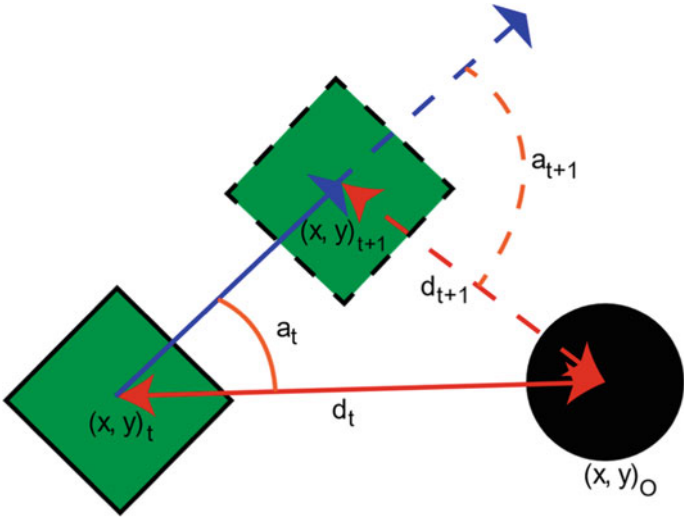


Fig. 4 Look-ahead computation

For the partition approach, the next state following a selected action was estimated by simply "rotating" the observed state seen in Fig. 3 according to the current action. This is shown in Fig. 5.

Next, the function to determine what action to take in any given state was defined as follows:

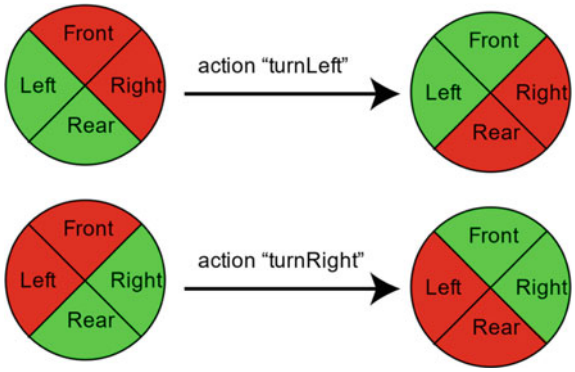


Fig. 5 Next-state estimation for cardinal partition approach

Algorithm 1

```
Pick a random decimal number n ∈ [0,1)
if n < r then
    return random action A ∈ A'
end if
Max value q ← −∞
A ← keepGoing
for each action a : A' do
    if Q(S_t, a) > q then
        A ← a, q ← Q(S_t, a)
    end if
end for
return A
```

Next, the estimated reward was calculated. For the relative angle and distance approach, the expected reward was based on the predicted distance to the nearest obstacle. This meant that if the agent was less than 10 pixels away from an obstacle, it counted as a collision, and the agent received a penalty in the form of a negative reward. Next, if the agent was closer than a "safe distance" of 20 pixels, it received a low, positive reward. Only, if this safety distance was exceeded, the agent would receive a larger reward. For the partition approach, the reward was based on the number of obstacles in the front partition after "rotation:"

Finally, combining Algorithm 1 with (1, 2, 3, 4) resulted in the following final algorithm used for agent learning:

Algorithm 2

1: Initialize $Q(s, a)$ with default values 0
2: Select initial action A
3: **while** true **do**
4: Get sensor readings d_t, a_t for current agent position t
5: Determine state S_t using (1) and (2)
6: Apply rotation associated with A to $\vec{V}$
7: Calculate d_{t+1} and a_{t+1} for next state S_{t+1} using (3) and (4)
8: Get action $A_{t+1} \in A'$ with the highest Q-value for S_{t+1} from $Q(s, a$
9: Calculate reward R based on d_{t+1}
10: $Q(S_t, A) \leftarrow Q(S_t, A) + \alpha \times (R + \gamma \times Q(S_{t+1}, A_{t+1}) - Q(S_t, A))$
11: $A \leftarrow A_{t+1}$
12: **if** A is not WAIT **then**
13: Translate agent position according to $\vec{V}(S_t \leftarrow S_{t+1})$

Table 1 Control variable values

Variable	Value
r	0.038
α	1.0
γ	0.8
W	25
H	20

The values of control parameters H, W, α, and γ used in the implementation are shown in Table 1.

3.3 Obstacle Generation

By default, a preset number of obstacles were generated and distributed in the environment for the agents to avoid. No explicit hard world border was defined. In addition to the randomly distributed obstacles, lines of obstacle objects were added, surrounding the environment. This was done in order to make agents learn to stay within the screen using the same algorithm as for obstacle avoidance. How this addition affected the agents' behavior will be discussed later. When agents did wander off screen, they were repositioned in their original starting location and continued exploring. In the simulator, agents were also considered obstacles. The sensors used could detect both static obstacles, and other agents present in the scene.

4 Results

A set of experiments were performed in order to examine the agents' ability to learn to avoid obstacles and deal with congestion. Firstly, it was of interest to test a single agent in a static environment, to see if the basic concept of the implementation worked as intended. A single agent was added to an environment of size 1184×761 pixels, with 250 obstacles randomly distributed. Both approaches to the state space discretization and mapping produced similar visual results:

In Fig. 6, we see a single agent with trail and status label avoiding obstacles. On the left, the agent was able to stay more than 20 pixels away from the obstacle when selecting action "KeepGoing," and received a reward based on distance. In the right-hand part of the figure, the agent came too close to an obstacle, and thus received a reward of -10. It is also possible to see from the trail that the agent was able to navigate and avoid the obstacles. Selected screenshots of single-agent navigation are presented in Fig. 7:

As previously mentioned, no explicit world border was defined in the reward function. To discourage agents from leaving the screen, a border made up of static

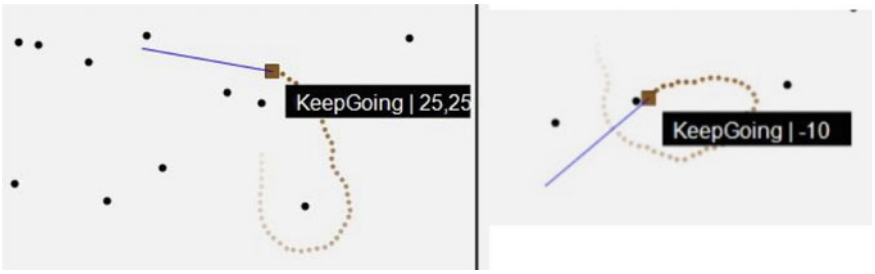


Fig. 6 Single agent with action and reward label

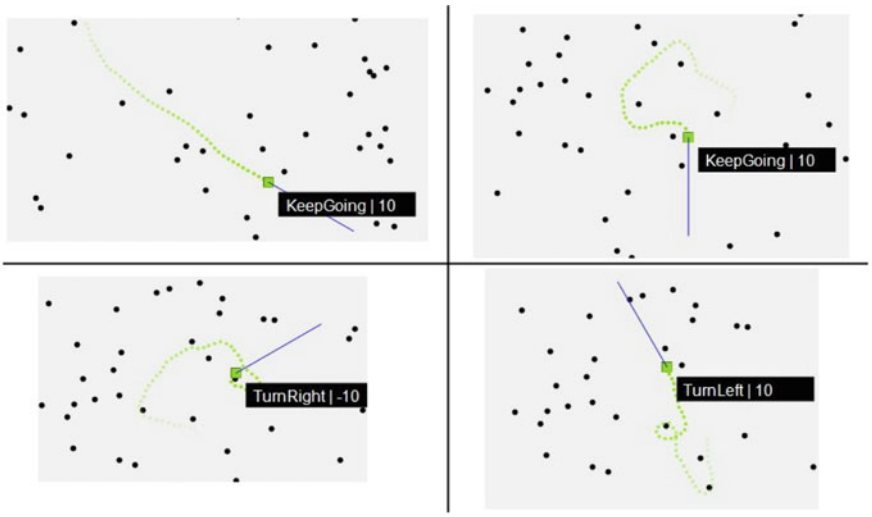


Fig. 7 Lime agent navigating

obstacle objects was added, in addition to the randomly distributed obstacles, as can be seen in Fig. 8. Agents would sometimes leave the scene even with this border in place but were more likely to swerve back into view, also seen in Fig. 8:

This border also served as a mechanism to introduce space restraints and congestion to the system. To get a better understanding how this border affected the agents, the number of times agents left the scene over time, with and without the border was recorded. 25 agents and 250 obstacles were added to the environment. Ten instances of the simulator were run simultaneously for 10 min, and their numbers averaged to account for the randomness inherit in the reinforcement learning approach.

Since this benchmark was based on time, the wait action was disabled. Otherwise, the agents would potentially get good results by simply standing still for the duration of the simulator.

Figure 9 shows that agents had left the scene more than 250 times after 10 min (600 s) when no border was defined. When introducing the obstacle border, agents

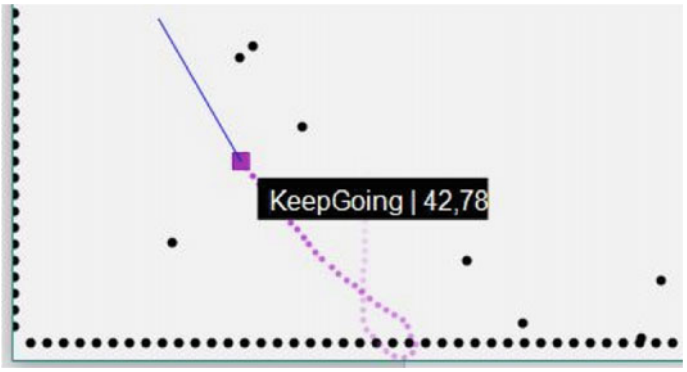


Fig. 8 Agent responding to border of obstacles

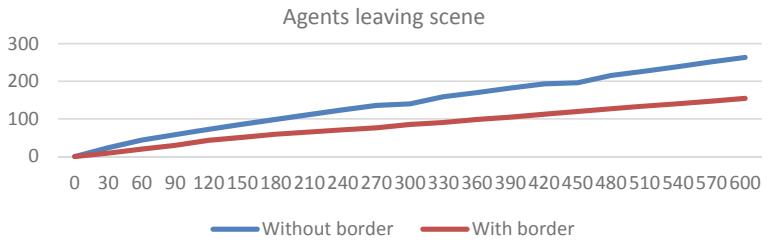


Fig. 9 Agents leaving scene with and without obstacle border

were more inclined to stay within the screen and had wandered off 154.4 times on average.

Another interesting aspect regarding AGVs in enclosed spaces is to examine how the size of the environment affects congestion. In the next experiments, the size of the simulator window was reduced, and the number of collisions for each configuration was measured. As in the previous experiment, the number of agents used was 25 and the number of obstacles 250. Each configuration was run 10 times for 300 s, and the resulting values averaged. The size of the window in each experiment is listed in Table 2.

A plot of the size of available area compared to the total number of collisions after 300 s is shown in Fig. 10.

Table 2 Simulator environment sizes used in experiments

Window size in pixels	Environment size in pixels	% of default area
1200 × 800 (default)	1184 × 761	100
900 × 600	884 × 761	55
600 × 400	584 × 361	23.4

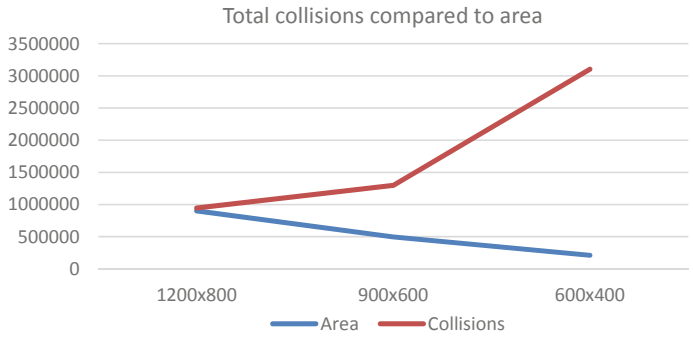


Fig. 10 Total collisions compared to simulator area

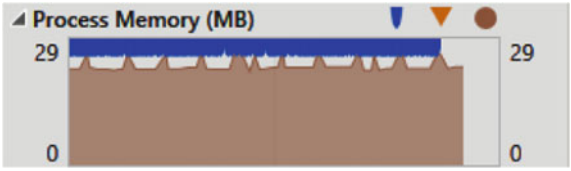


Fig. 11 Screenshot showing memory usage of the implementation

Finally, the performance of both approaches was measured and compared based on the number of collisions per distance in pixels traveled by each agent. For this experiment, the number of agents was 20; the number of obstacles was 200, and the wait action was enabled. The screen size was left at the default, and the simulator was run for 5 min in both configurations. For the relative angle/distance approach, the average number of collisions per distance in pixels was 0.337. For the cardinal partition approach, the resulting average was 0.072. Both approaches had a similar memory usage of less than 30 MB of RAM, as seen in Fig. 11.

5 Discussion of Results

As seen in Figs. 6, 7, and 8, agents were able to avoid obstacles and received a negative reward when they failed to do so. As seen in Fig. 8, the agent wandered off slightly, then was able to reorient itself and navigate back into view. This screenshot was picked specifically to demonstrate this behavior, as agents would sometimes disappear completely off screen. In those cases, the agent was automatically repositioned to its random starting position. No penalty was given for leaving the scene other than obstacle proximity. The effect of this is even clearer in Fig. 9, which compares the number of times agents left the scene with and without the border in place. From this, it is possible to see that the agents were incentivized to stay within the screen solely based on the learned obstacle avoidance strategy.

Next, the effects of limiting the space of the agents were examined. In these experiments, a collision was counted every time an agent received a negative reward. When looking at Fig. 10, we see the relationship between the number of collision and total area. As one might expect, there was an inverse correlation. What is interesting is that when reducing the area to 55% of the default, collisions only went up by 37%. However, this did not hold when reducing the environment further to 23%. In this case, collisions rose by 228%, as seen in Fig. 10. The limitation of agent actions is also expected to have had an impact, since agents could only turn left or right or keep going straight. With less space to "idle" in, constantly moving while still maintaining a safe distance becomes much more difficult due to congestion. This suggests that this approach only performs acceptably down to a certain environment size while maintaining the number of obstacles and agents. In larger environments, this becomes less of a problem, mostly due to limited agent vision. An agent does not consider an obstacle it cannot detect.

Furthermore, when comparing the performance of the two approaches, it appears that the cardinal partition approach outperforms the relative angle/distance approach when we only consider collisions per distance traveled. This is interesting since that approach only relied on 16 states, while the angle/distance approach used 500 states. The next-state calculation was also simpler and was achieved in practice by simply swapping the indices of the four partitions according to the selected action. Being able to achieve better of performance with a reduced number of states could be beneficial when running this system on low-power hardware for real-life applications and suggests that a partition-based approach shows some promise for further investigations.

5.1 Further Work and Recommendations

The introduction of a negotiating mechanism to help with space allocation in small environments is expected to benefit the system and help deal with congestion. In practice, all agents have to share a common good in the form of space, or more specifically, the maximum distance from obstacles. Negotiations could be in the form of agent-to-agent, or a mediator agent could be introduced, like in [7]. Furthermore, using multiple, or even other types of sensors, to define the state space should be looked into. Using LiDAR-only has been proven to be a drawback previously [11], and introducing visual information did improve performance in certain situations, as shown in [12].

Further expanding or enhancing the state space, as demonstrated in the cardinal partition approach, should also be looked into. Even though that approach showed promise with only 16 states, introducing additional granularity to the partitioning should also be investigated. Adding additional states would also mean that manually defining rules of action for each state would be increasingly difficult, strengthening

the argument that the agents should create their own models without human intervention. In summary, using a small, discretized state space as proposed did show some promise, but continuous state spaces and actions should also be investigated.

6 Conclusion

Two methods to simplify and model the state space for autonomous mobile agents were presented, and a simulator was implemented and demonstrated in software. The presented approaches demonstrated how a complex environment with multiple obstacles could be simplified using a discretized state space based on a single LiDAR-like sensor. Agents were implemented with an internal representation of their environment created in real time, and an on-policy reinforcement learning algorithm was used in order to learn to avoid randomly distributed obstacles. Discretizing the state space allowed for a simple, tabular approach to reinforcement learning, which is less resource intensive than high dimensional or even continuous state spaces.

Results showed some promise using a single LiDAR-like sensor to construct simple state spaces, which can be useful when dealing with cheap, low-power equipment in the real world. A complex environment with multiple agents and 200+ obstacles was broken down into a small state space, requiring less than 30 MB to run including the graphical user interface, while still being useful for obstacle avoidance. Using a similar approach when building and configuring RL systems for real-world AGVs may make simple proof-of-concepts easier to deploy rapidly. Further, tweaking to low-level actuator control can thus be made while still providing a consistent interface to end users and developers.

References

1. Back to the future? Airline sector poised for change post-COVID-19. https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/back-to-the-future-airline-sector-poised-for-change-post-covid-19. Last accessed 2021/06/08
2. Baggage IT Insights 2020. https://www.sita.aero/resources/surveys-reports/baggage-it-insights-2020. Last accessed 2021/03/04
3. Meyer G, Främling K, Holmström J (2009) Intelligent products: a survey. Comput Ind 60(3):137–148
4. Sella R, Rassõlkinb A, Wanga R, Otto T (2019) Integration of autonomous vehicles and Industry 4.0. In: Proceedings of the Estonian Academy of Sciences. Estonian Academy Publishers, Tallin, Estonia, pp 389–394
5. Thürer M, Fernandes NO, Stevenson M, Qu T, Huang GQ (2021) Self-organizing material flow control using smart products: an assessment by simulation. J Ind Prod Eng 38(2):148–156
6. Bussmann S, Schild K (2000) Self-organizing manufacturing control: an industrial application of agent technology. In: Proceedings fourth international conference on multiagent systems. IEEE, pp 87–94

7. Hallenborg K, Demazeau Y (2006) Dynamical control in large-scale material handling systems through agent technology. In: 2006 IEEE/WIC/ACM international conference on intelligent agent technology. IEEE, pp 637–645
8. Tumer K, Agogino A Distributed agent-based air traffic flow management. In: Proceedings of the 6th international joint conference on autonomous agents and multiagent systems. Association for Computing Machinery, New York, NY, pp 1–8
9. Duc DN, Huu TT, Nananukul N (2020) A dynamic route-planning system based on Industry 4.0 technology. Algorithms 13(12):308
10. Cheng Y, Wang GY (2018) Mobile robot navigation based on lidar. In: Chinese control and decision conference. IEEE, pp 1243–1246
11. Malavazi FBP, Guyonneau R, Fasquel J-B, Lagrange S, Mercier F (2018) LiDAR-only based navigation algorithm for an autonomous agricultural robot. Comput Electron Agric 154:71–79
12. Su Z, Zhou X, Cheng T, Zhang H, Xu B, Chen W (2018) Global localization of a mobile robot using lidar and visual features. In: 2017 IEEE international conference on robotics and biomimetics. IEEE, pp 2377–2383
13. Jansson AD (2021) Simulation of obstacle avoidance for multiple autonomous vehicles in a dynamic environment using Q-learning. Int J Comput Inf Eng 15:267–272
14. Sutton RS, Barto AG (2015) Reinforcement learning: an introduction, 2nd edn. The MIT Press, Cambridge, Massachusetts, MA

Attitudes Toward Time and Attitudes Toward Debt: Structural Equation Modeling Results

M. A. Gagarina, T. A. Nestik, and A. N. Nevryuev

Abstract The paper describes the results of an empirical study of relationships between time attitudes and debt attitudes. The total sample was 3022 respondents, aged 18–84. Inventories used: author's questionnaire for socio-demographic data, express debt behavior instrument, long-term orientation questionnaire, short version of the Zimbardo time perspective inventory, and a short Russian version of the consideration of future consequences scale. Future, hedonistic present and positive past TPs reinforce long-term orientation, while negative past can have different influences on its components. The motivational value and behavioral components of long-term orientation support a willingness to take into account future consequences, which in turn increases the willingness to lend and reduces the willingness to borrow and fulfill obligations. Statistics: Structural equation modeling was carried out with the Mplus 7 program. Results: It was found that different types of time perspectives are differently associated with long-term orientation, but long-term orientation is positively associated with consideration of future consequences. Debt behavior is related in different ways to consideration of immediate consequences: positively with avoidance of borrowing and negatively with debt rationality and disapproval of lending.

Keywords Time perspective · Long-term orientation · Consideration of future consequences · Debt attitudes · Structural equation modeling

M. A. Gagarina (✉) · A. N. Nevryuev
Financial University Under the Government of the Russian Federation, Moscow, Russia
e-mail: MGagarina@gmail.com

M. A. Gagarina · T. A. Nestik
Institute of Psychology, Russian Academy of Sciences, Moscow, Russia

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_43

1 Introduction

1.1 Present Study

This study was conducted during the COVID-19 pandemic. The COVID-19 pandemic has had a significant impact on the lifestyles of people around the world. The study of debt behavior is of particular interest, since it is this that is an indicator of changes taking place in society and reflects expectations about the future. As a working definition, the following will be used: "debt behavior is a type of economic behavior, including making decisions and actions to choose a credit institution (or an individual), borrowing and lending money, and fulfilling monetary obligations" [2] (pp. 51–52).

1.2 Structure of Debt Behavior

That is, debt behavior includes three aspects: (1) willingness to borrow, which indicates the need to meet the need immediately, using the resources that will be available in the future; (2) willingness to lend, which is a manifestation of one's own willingness to postpone or limit the satisfaction of needs in order to help another; (3) the fulfillment of debt obligations, reflecting the willingness to regulate the needs in the present in order to pay back for what was in the past. Debt behavior is determined by a number of factors, including both external (e.g., the situation in the country, cultural characteristics, availability of loans, etc.) and internal—psychological characteristics of an individual [12]. But our model is not about behavior, it is about mental representations of debt behavior, which will be designated below as attitudes.

1.3 Relationship Attitudes and Debt Behavior

At the same time, conflicting results have been accumulated in relation to the links between attitudes and debt behavior. Although some studies have not found a relationship between attitudes toward debt and debt behavior [11, 19, 20], most studies, including longitudinal ones, nevertheless indicate that attitudes are significant predictors of debt behavior and can be used to make predictions about behavior: for example, the presence of financial debt among university students is positively associated with greater debt tolerance [1]; the affective credit component is a predictor of purchasing behavior and actual use of credit cards [5, 6]. The ambiguity of the results is due to the fact that debt attitudes have a complex structure. Longitudinal research has shown that debt attitudes include at least two independent components: debt fear and debt utility, which are significant predictors of educational debt [4]. At the same time, a more detailed analysis of the work indicating the lack of connection [20] confirms

the above said only in the case of assessing debt attitudes with particular debt attitude scale [1], while the attitude to property and expenses is a significant predictor of debt. Thus, in the case of debt attitudes, the rule that was included in classical handbooks on social psychology applies: Attitudes reliably predict behavior under the condition that they are strong, relevant to the observed behavior, and the actions of other factors are minimized [14].

1.4 Predictors of Debt Attitudes

As socio-psychological predictors of debt attitudes, we consider time perspective [12–14], long-term orientation [15, 16], consideration of future consequences [7, 23]. Time perspective describes a person's idea of his past, present, and future, as well as endowing these periods with different meanings and significant [25]. Consideration of future consequences reflects "the extent to which people consider the potential distant outcomes" [23] (p. 743). Long-term orientation describes cognitive, affective, behavioral, and motivational value aspects of attitude toward distant future [18].

2 Methodology

2.1 The Aim of This Study and Hypothesis

The aim of this study is to find out the relationship between time perspective, long-term orientation, and consideration of future consequences and debt attitudes.

Hypothesis: The components of a balanced time perspective: a positive past, a hedonic present, and a future enhance long-term orientation, which increases the focus on consideration for distant consequences, which increases the willingness to borrow, lend, and fulfill debt obligations.

2.2 Methods

Author's questionnaire, including social demographic data was used. To measure debt attitudes, we used standardized express debt behavior instrument [3]. It includes 14 items that form three scales: debt-related rationality (α−0.7); avoidance of borrowing (α−0.8); disapproval of lending (α—is 0.7). Scales include items describing both attitudes and real actions.

Long-term orientation. To assess long-term orientation, we used standardized long-term orientation questionnaire (LTO) by Nestik [18]. The scale comprises 28 items, which are grouped into eight subscales: Interest in the long-term future

(α−0.8), optimism about the long-term future (α−0.7), value of a time-stable achievements and a trace in history (α−0.8), value of continuity and traditions (α−0.8), value of continuous self-development (α−0.8), belief in long-term planning utility (α−0.8), belief in rewarding long-term efforts (α−0.8), and long-term goal-setting and planning (α−0.8). Higher scores on the scales indicated a higher level of long-term orientation.

Time perspective. To assess the time perspective, we used short version of the Zimbardo time perspective inventory (ZTPI–S) [10]. The questionnaire comprising 15 items in total, three items for each scale: negative past, hedonistic and fatalistic present, future, and positive past. Higher scores on the scales indicated a higher level of time perspective.

Consideration of future consequences. A short Russian version of the consideration of future consequences scale (CFC-6) was used [17]. Only total scale consideration of future consequences was included into analysis, with higher scores indicating a greater orientation to the future.

2.3 Sample

Representative Russian sample involved 3022 respondents from all federal districts: *M*—1531 (50.7%), *F*—1491 (49.3%). Age from 18 to 84 years old, $M = 40.51$ and $SD = 10.52$.

To analyze the direction of the relationship between the various components of long-term orientation, time perspective, consideration of future consequences and debt attitudes, a model was built that included the variables described above. Before modeling, a correlation analysis was performed to confirm the presence of significant relationships between the variables (not presented in the article). For correlational analysis, we used SPSS 23.0. Structural equation modeling was carried out with the Mplus 7 program (Muthén and Muthén, 1998–2012).

3 Results

3.1 Model Description

During the analysis, a model which included 14 elements, was tested. Elements were divided into four blocks: The first block is time perspectives, the second block is scales of long-term orientation, the third block is consideration of future consequences, and the fourth block is debt attitudes. The connections were built in the following way: The attitude toward one's past, present, and future determines the long-term orientation, leading to a different willingness to take into account future consequences, which is associated with debt attitudes.

3.2 Hypothesis Testing

The hypotheses were refined according to the hierarchy of relationships. In accordance with hypothesis, 1 "indicators of TPs—negative past and fatalistic present reduce the long-term orientation and consideration of future consequences and are associated with a high level of borrowing and low level of fulfillment of obligations" and hypothesis; 2 "positive past, hedonistic present and future enhance long-term orientation and consideration of future consequences and increase willingness to lend", we checked links between TPs, long-term orientation, CFC, and willingness to borrow, lend, and fulfill obligations.

3.3 Test Results

To test our hypotheses, we used structural equation modeling, the maximum likelihood estimator. The results of testing the model (see Fig. 1) demonstrated an acceptable level of compliance with the initial data: $x^2 = 141.919$, df = 455 $p < 0.001$, CFI = 0.976, TLI = 0.948, RMSEA = 0.023, 90 CI [0.018, 0.028], and SRMR = 0.016.

Debt-related rationality and disapproval of lending are reduced and avoidance of borrowing is enhanced by consideration of future consequences, which is directly dependent on the components of long-term orientation: value of continuity and traditions, value of continues self-development, and long-term goal-setting and planning. Orientation toward the past, both positive and negative, strengthens value of continuity and traditions; negative past TP weakens, and positive past TP strengthens

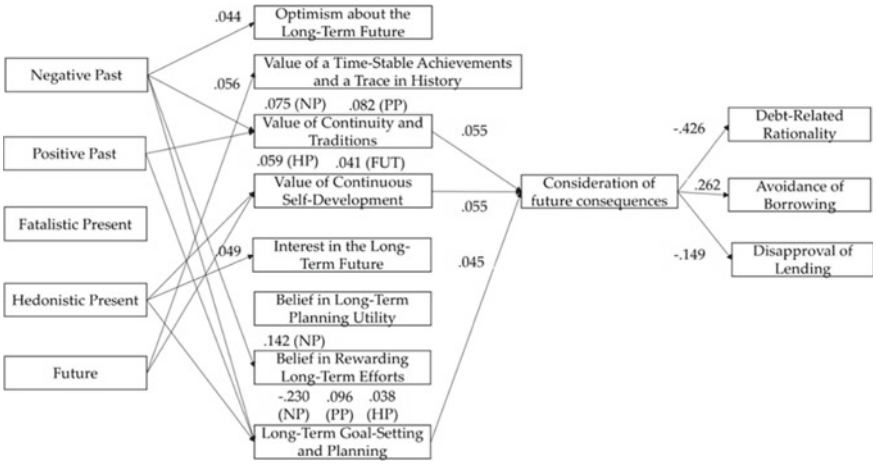


Fig. 1 Relationship between time perspectives, long-term orientation, consideration of future consequences, and debt attitudes. Standardized regression coefficients are presented

long-term goal-setting and planning, and future and hedonic present TPs strengthen value of continuous Development.

4 Discussion

4.1 Relationship Between Time Perspective and Long-Term Orientation

Future, hedonistic present and positive past TPs reinforce long-term orientation, while negative past can have different influences on its components. The value and behavioral components of long-term orientation support a willingness to take into account future consequences, which in turn increases the willingness to lend and reduces the willingness to borrow and fulfill obligations. The links between long-term orientation and attitudes toward borrowing and lending look quite understandable: A person oriented toward long-term consequences is ready to lend to other people, thereby investing in his social capital [22], and tries not to borrow, since such behavior is assessed as irrational in the Russians' mentality [2].

4.2 Relationship Between Consideration of Future Consequences and Debt Attitudes

Why, then, an increased focus on future consequences leads to a decrease in the willingness to make efforts and plan steps in obtaining and paying off loans? Apparently, in this case, a negative attitude toward loans is reflected: Selectivity in the choice of credit conditions and consistency in debt repayment are viewed in the long term not as rational, but, on the contrary, as irrational and dangerous actions. In addition, it is likely that there can be an identification of debt as an obligation and debt as a loan [2]. In other words, the fulfillment of debt obligations is considered as a consequence of an error made when applying for a loan, and therefore is assessed negatively.

5 Conclusion

This study made it possible to establish that people who value the past, strive for self-development and plan for the future are guided by future consequences. In addition, focusing on future consequences increases the avoidance of borrowing and decreases debt rationality and disapproval of lending.

The limitations of this study include the fact that it was conducted during the COVID-19 period. In further studies, it is necessary to compare the obtained data

with "post-Covid" results. Such a comparison will make it possible to understand the specifics of attitudes toward time and debt in its dynamic.

Acknowledgements This research was funded by the Russian Foundation for Basic Research grant number 19-29-07463mk.

References

1. Davies E, Lea SEG (1995) Student attitudes to student debt. J Econ Psychol 16:663–679. https://doi.org/10.1016/0167-4870(96)80014-6
2. Gagarina MA (2021) Debt mentality of a person: an economic and psychological study. IPRASMoscow, p 275
3. Gagarina MA, Padun MA (2021) Debt behavior express inventory: development and validation. Psikhologicheskie Issledovaniya, 78, 3. URL: http://psystudy.ru. (In Russian, Abstr. In English)
4. Haultain S, Kemp S, Chernyshenko OS (2010) The structure of attitudes to student debt. J Econ Psychol 31:322–330. https://doi.org/10.1016/j.joep.2010.01.003
5. Hayhoe CR, Leach L, Turner PR (1999) Discriminating the number of credit cards held by college students using credit and money attitudes. J Econ Psychol 20:643–656. https://doi.org/10.1016/S0167-4870(99)00028-8
6. Hayhoe CR, Leach LJ, Turner PR, Bruin MJ, Lawrence FC (2000) Differences in spending habits and credit use of college students. J Consum Aff 34:113–133. https://doi.org/10.1111/j.1745-6606.2000.tb00087.x
7. Joireman J, King S (2016) Individual differences in the consideration of future and (more) immediate consequences: a review and directions for future research. Soc Pers Psychol Compass 10:313–326. https://doi.org/10.1111/spc3.12252
8. Joireman J, Strathman A, Balliet D (2006) Considering future consequences: an integrative model. In Judgments over time: The interplay of thoughts, feelings, and behaviors; L. Sanna & E. Chan; Oxford: Oxford University Press
9. Klicperová-Baker, M., Košál, J., Vinopal, J.: Time Perspective in Consumer Behavior. In Time perspective theory; review, research and application: Essays in honor of Philip G. Zimbardo; Stolarski M., Fieulaine N., van Beek W.; Springer International Publishing, Cham, Switzerland (2015). https://doi.org/10.1007/978-3-319-07368-2_23
10. Kostal, J., Klicperova-Baker, M., Lukavska, Jiri Lukavsky, K.: Short version of the Zimbardo Time Perspective Inventory (ZTPI–short) with and without the Future-Negative scale, verified on nationally representative samples. Time & Society, 25, 169–192 (2016). doi:https://doi.org/10.1177/0961463X15577254
11. Lachance M (2012) Young adults' attitudes towards credit. Int J Consum Stud 36(5):539–548. https://doi.org/10.1111/j.1470-6431.2012.01119.x
12. Lea SEG, Webley P, Walker CM (1995) Psychological Factors in Consumer Debt: Money Management, Economic Socialization, and Credit Use. J Econ Psychol 16:681–701. https://doi.org/10.1016/0167-4870(95)00013-4
13. Leonard S, Zhang JW, Howell R (2019) Spending well: how time perspectives impact consumer values and financial decisions among middle-aged adults. Res Hum Dev 16:135–155. https://doi.org/10.1080/15427609.2019.1670568
14. Mayers, D. Sotsial'naya psikhologiya. SPb.: Piter. 794 (2010). (In Russ.)
15. Nestik, T.A.: The socio-psychological mechanisms of long-term orientation. Institute of psychology Russian Academy of Sciences. Social and economic psychology, 4, 16–60 (2016). http://soc-econom-psychology.ru/cntnt/bloks/dop-menu/archive/g16/t1-4/s16-4-02.html (In Russian, Abstr. In English)

16. Nestik, T.A.: Long-term orientation of the person: the state and prospects of research. Institute of psychology Russian Academy of Sciences. Social and economic psychology, 3, 110–140 (2020). DOI: https://doi.org/10.38098/ipran.sep.2020.19.3.005.
17. Nestik, T.A.: Russian validation of the consideration of future consequences scale (CFC-14). Institute of psychology of the Russian Academy of Sciences. Organizational psychology and labor psychology, 3, 55–88 (2020). DOI: https://doi.org/10.38098/ipran.opwp.2020.16.3.003 (In Russian, Abstr. In English)
18. Nestik, T.A.: Socio-psychological predictors and types of personal long-term orientation. Psikhologicheskii Zhurnal, 4, 28–39 (2021). (In Russian, Abstr. In English).
19. Norvilitis JM (2014) Changes over Time in College Student Credit Card Attitudes and Debt: Evidence from One Campus. J Consum Aff 48(3):634–647. https://doi.org/10.1111/joca.12043
20. Norvilitis JM, Merwin MM, Osberg TM, Roehling PV, Young P, Kamas MM (2006) Personality factors, money attitudes, financial knowledge, and credit-card debt in college students. J Appl Soc Psychol 36(6):1395–1413. https://doi.org/10.1111/j.0021-9029.2006.00065.x
21. Sekścińska K, Rudzinska-Wojciechowska J, Maison D (2018) Individual differences in time perspectives and risky financial choices. Personality Individ Differ 120:118–126. https://doi.org/10.1016/j.paid.2017.08.038
22. Shoji M, Aoyagi K, Kasahara RJ, Sawada Y, Ueyama M (2012) Social capital formation and credit access: Evidence from Sri Lanka. World Dev 40:2522–2536
23. Strathman A, Gleicher F, Boninger DS, Edwards CS (1994) The consideration of future consequences: Weighing immediate and distant outcomes of behavior. J Pers Soc Psychol 66:742–752
24. Unger, A., Lyu, H., Zimbardo, P.G.: How Compulsive Buying Is Influenced by Time Perspective—Cross-Cultural Evidence from Germany, Ukraine, and China. International Journal of Mental Health and Addiction, 16, 525–544 (2018). https://ezpro.fa.ru:2696/10.1007/s11469-018-9942-4
25. Zimbardo, P.G., Boyd, J.N.: Putting time in perspective: A valid, reliable individual-differences metric. In Time perspective theory; review, research and application: Essays in honor of Philip G. Zimbardo; Stolarski M., Fieulaine N., van Beek W.; Springer International Publishing, Cham, Switzerland (2015). https://doi.org/10.1007/978-3-319-07368-2_2

Estimation of Programming Understanding by Time Series Analysis of Code Puzzles

Hiroki Ito, Hiromitsu Shimakawa, and Fumiko Harada

Abstract In programming education, it is desirable for instructors to stand beside learners and monitor their answering process to assess the individual's actual ability. However, it seems to be impossible in large-group lectures at educational institutions or newcomer education at companies. Therefore, instructors attempt to grasp the understanding status of many learners at once by using written tests and e-learning to find out the learners who need instruction. They examine the learner's knowledge such as algorithm and syntax. However, in reality, not a few learners fail to acquire the skill of writing source codes. This kind of situation implies that the programming ability of learners cannot be measured only by knowledge tests or the data obtained from answer results. The purpose of this study is to estimate the understanding of programming, focusing on the thinking process. This paper analyzes a time series of operations of learners working on code puzzles, where they arrange code fragments. Since we assumed learners with low understanding are different from those with high in terms of the consistency of blocks of code fragments to be touched, we modeled it using a hidden Markov model. The proposed method estimates their perspectives on how fragments are built up to achieve given requirements. The results of an experiment have shown that the calculated hidden Markov model produces meaningful interpretable values. Furthermore, the values show significant indices that machine learning models can explain the understanding of learners.

Keywords Programming education · Learning analytics · Computational thinking · Process oriented · Code puzzle · Time series · Remote tutoring

H. Ito (✉)
Graduate School of Information Science and Engineering, Ritsumeikan University, Shiga, Japan
e-mail: hirokiito6900@de.is.ritsumei.ac.jp

H. Shimakawa
College of Information Science and Engineering, Ritsumeikan University, Shiga, Japan
e-mail: simakawa@cs.ritsumei.ac.jp

F. Harada
Connect Dot Ltd., Tokyo, Japan
e-mail: harada@de.is.ritsumei.ac.jp

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_44

1 Introduction

According to the modern research trends in the field of programming education, many studies try to predict the performance and the learners who drop out [1]. They examine the knowledge of the learners. Many students cannot create correct sources although they pass written tests that ask for their knowledge on grammar and how to express algorithms. As a result, they drop out of the programming course [2, 3]. Nikula et al. [2] mentioned that one of the factors is motivation, but also that there is a clear correlation with course scores. This means that we need to properly assess students' programming abilities and also provide them with appropriate interventions.

The programming ability cannot be measured only with the learner's knowledge. This is because programming skill also requires the ability to construct program elements logically with a perspective [4, 5]. Nesbit et al. [6] also speculated that intelligent tutoring systems (ITSs) within Computer Science and Software Engineering Education contain a higher proportion of procedural learning goals while a lower proportion of conceptual learning goals. Programming thinking ability is the one to assemble the components of a program with a perspective, given requirements to be satisfied. Any method has not been established yet to measure programming thinking ability. Previous researchers have considered that the backgrounds, prior knowledge, cognitive abilities, working time, and learning attitudes may be related to the performance of learners [7]. However, these information do not allow us to provide appropriate interventions to students. Some studies have tried to estimate learners' programming ability from other abilities. However, most of the existing educational institutes use measuring methods biased into knowledge aspects because of their easiness in assessment. As the result, many learners fail to understand the intention of the programming task, which prevents them from acquiring the ability to realize it. There is a strong demand for a method to grasp the students' understanding chronologically, taking their programming thinking ability into account.

This study analyzes a time series of operations of a learner struggling with a code puzzle, in which the learner arranges code fragments. The proposed method aims to estimate their perspectives on how fragments are organized to achieve given requirements. In addition, to evaluate whether the estimated perspective status is valid, this paper uses machine learning to verify that the labels given by the instructor can be explained. A trained model can quantify the learner's level of understanding. It makes it easier to detect those who need additional instruction.

2 Programming Education Support

2.1 Current Programming Education Support

This paper discusses understanding analysis that focuses on the thought process of programming. According to the schema theory in cognitive psychology [8], when

humans solve a problem, they read the problem to create a perspective, a way to solve it. That is, in programming education, without assessing how to create a way to solve the task, it cannot be said that logical thinking ability is being assessed. In the evaluation of programming ability, it is essential to establish a learning support method that focuses on the composing aspect rather than the knowledge aspect.

2.2 Programming Learning Through Code Puzzles

This research uses a code puzzle as a platform for learning. A code puzzle is a programming task where the learner rearranges code fragments such as source code and pseudocode to assemble them to achieve a specified requirement. The code puzzle is inspired by Programming Puzzle proposed by Parson and Haden [9] Scratch developed by MIT is a famous example that takes the approach. Parson et al. insisted that code puzzles are more effective for beginners because they work better at nurturing logical thinking than full coding. Moreover, code puzzles request learners to make the logic flow, unlike the blank-filling problem. It is an exercise format suitable for developing computational thinking, which has been emphasized recently. In addition, code puzzles facilitate to obtain features of the learner's learning behavior because the learner takes actions of selecting, moving, and placing blocks in order. They enable us to analyze how the learner builds perspectives to achieve the programming task.

2.3 Related Works

According to a recent report on intelligent tutoring systems [10], adaptive feedback has been getting popular in recent years. They give feedback in a variety of ways. The major difference is adaptive feedback gives step-based hints, while conventionals give summative feedback on the submitted code. It implies the inherent difficulty in developing an understanding of the structure of the program and the associated problems. The method proposed by Jadud [11] is an early study that identifies learners who need guidance using compilation errors. However, it focuses only on the error. It is not possible to measure why the error occurred and how deeply the learner understands. Mysore and Guo [12] have proposed Porta, a Web system that can identify the part where the learner is struggling. Porta calculates areas difficult for a learner based on the attention to the material but does not calculate an individual's level of understanding. Since it does not judge understanding failure, it cannot detect where misunderstandings exist. Guo [13] has proposed an interface that supports one-to-many programming learning on the spot. They presented its implementation, Codeopticon. It is a tool that enables constant monitoring of modifications of learners' codes, but it does not provide the instructor with an intuitive grasp of whether the learner is struggling without perspective. Asai et al. [14] identified the cognitive load on learners and the factors that causes the cognitive load by the blank-

filling assignment. However, the blank-filling assignment hinders building the flow of logic. It cannot be said that it considers programming thinking. Many of these existing researches focus only on the aspect of knowledge. They do not estimate the understanding based on programming thinking. As a study focusing on behavior, Ihantola et al. [15] have estimated the difficulty level of a task using a decision tree from the answering process such as the answering time and the keystroke of a programming task. They suggest that the answer process brings a significant difference in learner's understanding. However, they did not mention programming thinking ability. Their method cannot estimate its factors. Another process-oriented approach is sometimes referred [16] to as WATWIN score proposed by Watson et al. [17, 18] and EQ score proposed by Jadud [11] and improved by Tabanao et al. [19]. Both of them are based on the compilation behavior. The approach is mainly to evaluate whether the type of compilation error has been solved. These approaches seem to be good ideas based on our experience. However, we also know from experience that compilation behavior and work time do not always represent the learner's ability. Ito et al. worked on estimating understanding from the process of working on code puzzles. The work succeeds to estimate 80% of labels given by the instructors to show learner's understanding. However, to estimate perspective based on composing aspects of programming, it is necessary to analyze a time series of operations by learners working on the task.

3 Educational Support Using the Behavior of Code Puzzles

3.1 Method of Estimating Understanding Based on Operation Process

The study aims at estimating the understanding from how good perspectives learners have in the process of tackling a task, rather than from the results of the task tackling as in conventional educational support systems. Figure 1 shows a schematic diagram of the proposed method. First, the learner solves a certain task in the form of a code puzzle introduced in Chapter "Weight-Based Dynamic Hybrid Recommendation System for Web Application Content". The interface used for the code puzzle is a proprietary application that runs on the Web. This application obtains the learner's operation log. The behavior of these operation logs should be different depending on the level of understanding of each learner. For example, those who cannot grasp the role of each block would not be able to behave consistently in the operation to build the block. On the other hand, those who know the role each of the blocks has, but do not know how to assemble them would spend more time arranging the blocks in the right order, even if their behavior is consistent. Their operation logs would represent puzzlement. Based on this hypothesis, an educational support system is developed so that it enables us to determine a learner's understanding based on the history of performed operations. The method uses a state-space model to describe

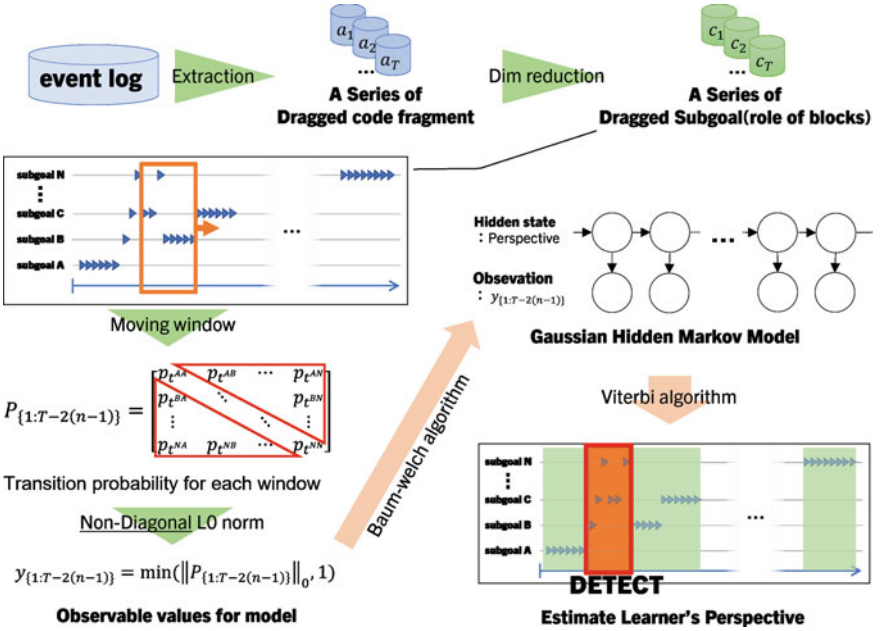


Fig. 1 Schematic diagram of the method

the learner's process of working on the problem, as shown in Fig. 1. It founds on the idea that the learner's process of programming engagement may have a hidden state of understanding at each point in time, behind the observable behavior of the learner. Observable behavior means that the learner moves a block. The method estimates the hidden states from the created state-space model. Eventually, individual feedback is given to the learner, according to the estimated hidden states.

This method trains a state-space model based on the observed sequence of blocks the learner moves. To complete a whole program, learners must achieve many subgoals. Learners who understand programming well move blocks with perspectives to achieve specific subgoals. From the hidden state of the state-space model, the method examines whether there is a significant difference between learners with perspectives and those without perspectives. If there is a significant difference, instructors can use that information to determine the understanding levels of each individual. Suppose the hidden Markov model's output can predict the understanding level the instructor judges. Such a predicted hidden state is significant. The predicted hidden state enables instructors to find new feedback, which they could discover only when they keep watching learner's behavior in the traditional results-based methods. When this model is applied to a new learner, more than one prediction result on the understanding level may be produced. They are sorted with the classification probability. The result of the highest probability is chosen so that the learner who needs instruction should be found immediately.

3.2 Collecting Learner Behavior by Code Puzzles

The method proposed in this study uses the tools shown in Figs. 2 and 3 to collect learner's characteristic behaviors to measure the understanding. In this research, the notation of the program is based on PAD proposed by Futamura et al. [20]. In the proposed method, an exemplary program or source code that satisfies all the requirements given in the task is divided into code fragments or pseudocode with functional cohesion. A code fragment generated by the division is referred to as a block in this paper. As shown in Fig. 2, blocks for assembling the program are displayed for a task. As shown in Fig. 3, the learner assembles the program using the blocks so that all the given requirements are satisfied. The proposed method examines the learner's thinking processes during solving the code puzzle to judge whether the learners have acquired programming thinking or not. For example, a person who cannot grasp the roles of the blocks will constantly touch completely misplaced blocks. In this case, the behavior of the operation will be inconsistent. Others who know the roles of the blocks but do not know how to assemble them will spend more time than necessary. Their behavior expresses that they get lost, even if they touch consistent blocks. It is assumed that the thought process in answering the task is characterized by the behavior of moving the blocks, depending on the understanding of the learner. The learner can switch between the task sentence screen and the drawing screen from the tab at the top of the tool. The learners enter his or her name; the assignment text appears on the screen, and when they interpret the assignment, they move to the drawing screen to initiate working on the programming. If they have any hesitation during the process, they will check the assignment text again, and when the program is completed to some extent, they will examine their own program in the execution and testing screen. The reason for separating the screens by tabs is to collect data on the time the learner spends interpreting the subject of the assignment, creating the program, and checking and testing the source code. The learner finishes solving the task by pressing the submit button, when they complete the task.

The tool collects a discrete value indicating which block is dragged. However, even small programs used for educational purposes generally consist of at least 20 lines of code fragments. This exceeds patterns that can be handled by the hidden Markov model for the amount of data collection assumed here. Therefore, this study divides the code fragments used in the program into subgoals that are assumed in the programming assignment and gives them the same label in advance. The labels are then used as observation variables. In other words, subgoals are key points included in an assignment that the learner should attain.

3.3 Grasping the Perspective Status of Each Learner

Understanding should not be treated in a single aspect. An assignment contains multiple subgoals. The level of understanding of each subgoal changes over the time

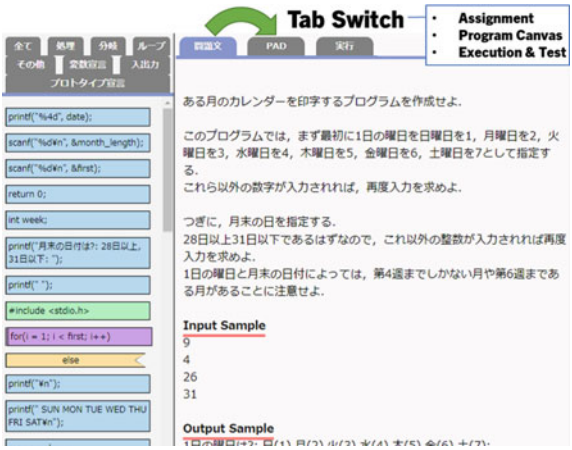


Fig. 2 Question text

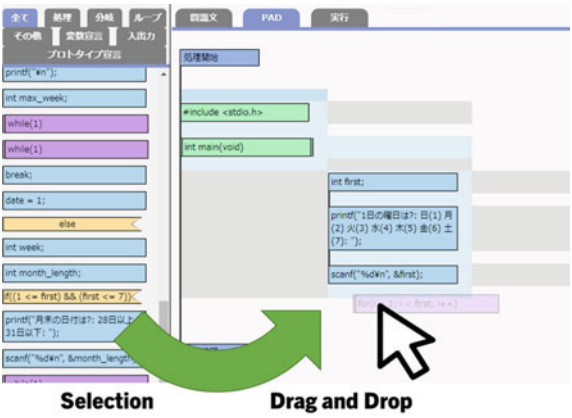


Fig. 3 Coding canvas

spent on the assignment. It is inevitable to observe the learner's process for instructors to understand it. In this method, the understanding of learners is estimated using a time series. In other words, the understanding depends on each point at which an operation is performed. Therefore, each operation is an observed data while the understanding at that time point to be in that hidden state. Now, since both of the observed data and the hidden state are discrete, the hidden Markov model is applied to represent the state-space model.

Let us now explain the method of calculating consistency using the concept of subgoals. We have explained this study divides the code fragments used in the program into subgoals that are assumed in the programming assignment and gives them the same label in advance. The labels are then used as observation variables. For example, if a programming task has a subgoal to input a variable, the subgoal may contain several code fragments such as a message to prompt for input, a function to accept inputs, and error handling. In this method, these code fragments will be given

the same label so that they will be treated as the same group. Next, the transition probabilities are calculated for this group using a moving window. The values of the diagonal components of the transition probabilities are particularly high when a certain subgoal is consistently focused by a learner. In this way, if a learner can focus on one subgoal consistently, he or she is regarded to have a good perspective on that subgoal, which can be expressed by the behavior transitions. Here, this method calculates the norm of the non-diagonal component to summarize it into a single variable that represents whether or not an operation is being performed consistently. In this case, the norm is the L0 norm, not the general norm. That is, the L0 norm is the frequency that different subgoals are touched in the moving window. When the norm is large, the variable shows a high percentage of manipulation of various subgoals that is, it indicates that the learner is straying. However, the transition of a subgoal only once in a moving window can be considered as a moment where one subgoal has been conquered and the next subgoal is being started, rather than being lost. Therefore, single transitions are ignored. Finally, the method obtains which subgoals have been dragged. The observed variable is the norm of the non-diagonal component of the transition probability between the specified windows. This means the degree to which the learner is confused. This observed variable is the input to the hidden Markov model. That is, the frequency of operation of each subgoal is used as an observation variable, and parameters such as transition probabilities in the hidden Markov model are obtained by the Baum–Welch algorithm. Since the frequency of operation of each subgoal is used as an observation variable, it is expected the hidden Markov model can be divided into several hidden states depending on how the frequency of this operation changes. Then, the Viterbi algorithm is used to estimate the hidden state at each time point for each individual. There are three hidden states in the hidden Markov model in this study that is, three levels of visibility to the subgoal. The first is a state with no prospect at all; the second is a state with a prospect, and the third is a transitional state between them.

The hidden state for each estimated by the Viterbi algorithm should differ depending on the existence of the perspective. A learner who has no perspective on a subgoal will frequently experience a hidden state of lack of perspective when operating on the corresponding subgoal. On the other hand, a learner who has a perspective on a subgoal will frequently experience the hidden state of having a perspective at the time he or she is operating on the corresponding subgoal. There may be people who have no perspective at first, but who realize the role of the module in the process and gain perspective. For such people, a hidden state, which is a transitional state, may appear in the process, and they may switch to the hidden state of having a perspective.

3.4 Feedback for Each Learner

This method can find subgoals that the learners fail to have a perspective on by estimating their understanding from the solution process. It is never achieved with a system that grasps the learner's understanding based on the results of answers. This

method provides instructors an intuitive understanding of the learner's approach to the programming task from the list of operated blocks and the hidden states estimated with the hidden Markov model. This makes it easy to identify the areas that need to be focused on for learners who need guidance.

4 Experiment

4.1 Experimental Methods and Objectives

The purpose of this experiment is to clarify the proposed method can determine whether a learner has a perspective on a particular subgoal based on the learner's operation logs while working on the code puzzle. The subjects were 15 university students with various levels of programming knowledge. There is no time limit to acquire the process where learners are confused. Some subjects completed the task in 20 min, while others in 1 h and 30 min. Only, 8 achieved a program that produced the correct output in the end. Since there is no time limit, it is not always possible to say that the learners have a high level of understanding, even if they can write a program that behaves correctly in the end. The subjects must be able to use the code puzzle application without any problem. Before tackling the task, they were given a tutorial and worked on a practice question, to get used to it. All experiments were conducted online using the Web.

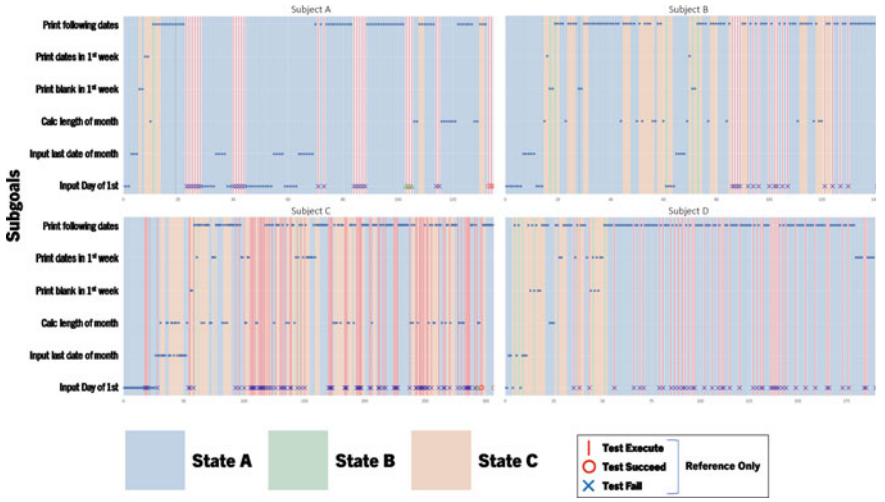


Fig. 4 Estimation results of hidden states by hidden Markov model

4.2 Hidden-State Estimation Results Using Hidden Markov Models

Applying the method described in Chapter "Measuring the Success of the Ngawi District Government Web Site Using the Delone and Mclean Model", the internal state of each learner is estimated with a hidden Markov model, which is shown in Fig. 4. Each plot represents a subgoal operated at a specific time point, while the color of the background represents the hidden state at that time point. In the figure, hidden state A appears more often when each subgoal is operated consistently, while hidden state C appears more often when inconsistent operations are performed on each subgoal. This method ignores subgoal transitions that occur only once that is, hidden states B and C will appear if two or more subgoal transitions occur during the specified time window. This information can be represented as shown in Fig. 4 so that we can quickly find out when the learner became confused, namely has lost perspective. It enables instructors or systems to notify learners of what they should think about next in an online manner.

5 Explanation of Understanding by Perspective Status

5.1 The Method to Explain Understanding Using Machine Learning

As shown in Sect. 4.2, estimating the internal state with a hidden Markov model, the method determines whether the learner took operations with perspective to achieve a certain subgoal. However, it has turned out those who take operations with a perspective do not always understand programming deeply. For example, let us watch the operation on the subgoal at the graph of subject D and the hidden state at that time points. We intuitively recognize that the hidden state A is continuing; however, the total number of operations is large. It can be said they got confused there. This is because they took operations under a perspective, due to which they got in trouble. In other words, even if they have a perspective, it may not always be the correct perspective. One of such subjects seemed to try to nest several layers of conditional and loop statements, which may cause a failure to achieve the right answer. Indeed, a perspective on a certain subgoal is considered to increase the rating of understanding of the learner. On the other hand, learners who have a state of perspective but do not know how to assemble it are considered to have a lower rating of understanding. It can be expected that the understanding of the code puzzle can be explained by the existence of perspective for each subgoal and the number of times it appears. This study uses machine learning to prove whether these features can correctly predict the understanding labeled by actual instructors. The labeling was conducted individually by four people who have been involved in university programming exercise classes

for more than three years. The final label is determined by a majority vote. Since the model is considered to be nonlinear, we use the random forest classification. As the number of data is relatively small, the generalization performance of the model is confirmed by leave-one-out validation.

5.2 Results of Estimation of Understanding

Table 1 shows the results of the experiments in Sect. 5.1.

The result of the leave-one-out validation was 1.0 in F_1-score. Since the result is skeptical, the five-part cross-validation is repeated ten times with different random states, which produced 99.3% in F_1-score. Since we completely separated the test data from validation data, the result shows the generalization performance correctly. This result shows the prediction result is almost perfectly coincident with the actual understanding. This means the perspectives can explain the learner's understanding and that the proposed indicator as internal state is trustworthy. Referring to the classification probabilities, the learners can be sorted in the order of their level of understanding. This makes it easier to find learners who need instruction.

Table 1 Results of the Experiments

user_id	Classification probability		Classification result	
	Low	High	True	Pred
1622345108	0.142	0.858	1	1
1622169741	0.234	0.766	1	1
1622469590	0.271	0.729	1	1
1621841538	0.314	0.686	1	1
1622434821	0.343	0.657	1	1
1631595816	0.448	0.553	1	1
1631600152	0.632	0.368	0	0
1631602622	0.652	0.348	0	0
1631605420	0.702	0.298	0	0
1631586473	0.764	0.236	0	0
1622145530	0.778	0.222	0	0
1622899809	0.801	0.199	0	0
1622430288	0.811	0.189	0	0
1631606028	0.827	0.173	0	0
1622262196	0.923	0.077	0	0

6 Consideration

6.1 Usefulness in Educational Settings

The results of this experiment show that it is possible to estimate the existence of perspective on the subgoals included in the programming task from the learner's operation logs when solving the code puzzle. In addition, the estimated perspectives of each subgoals were able to explain the instructor-labeled understanding. It can be said the existence of perspectives and the number of their occurrences were useful indicators This kind of time series analysis with an emphasis on the behavior of the answering process is essential for developing computational thinking. It is difficult for the instructor to grasp the understanding level of all learners. Especially nowadays, there are more and more opportunities to give lectures remotely. It is not always possible for instructors to directly watch over their students. In such a situation, if there is a system that can collect and list learners' learning behaviors online like this method, it is easy to provide feedback for the learners because it has strong possibilities to automate the evaluation for the feedback. In addition, information on the level of attention to each block can reveal which code fragments brought the learner difficulties in handling. Because of time series analysis, it is possible to grasp the learner's perspective status on the spot. When the system notifies the instructor or another system about learner confusion based on the perspective state, it can respond quickly to learners who are stuck in programming learning and prevent students from dropping out. Furthermore, it is possible to arrange learner's understanding levels in continuous values by estimating learners' perspective status using variables that focus on the answering process. It would enable education providers to use their human resources as much as they can to rescue learners who need help.

6.2 Issues and the Future

The goal of the future method is to be able to give a kind of summary sheet to learners, almost without relying on the instructor. We are working on making it possible to estimate understanding for each subgoal, rather than for each answer. This allows calculating the level of understanding for each subgoal so that such as hints or explanations of the subgoals that were not fully understood can be given to all students automatically. Also, we aim to estimate this understanding with models that can perform lookback, such as RNN and LSTM. Since there were relatively only 15 subjects in the evaluation of the understanding estimation model in this study, we must become skeptical that the model will perform as well as the paper in practical use. We will continue to conduct more subjects and evaluate the results reliably. Blikstein et al. [21] suggested that the understanding of learners can be measured using the clustering method with programming structure. Therefore, in our research, there is a possibility that the understanding and the factor of misunderstanding can be estimated with the clustering method.

7 Conclusion

This study proposes a method for estimating whether learners take operations with perspective to work on programming assignments by analyzing the time series of their answering process using code puzzles. As a result of the experiment, we were able to represent the learner's perspective as an internal state of the hidden Markov model. The experiment also reveals that a perspective does not always lead to the right answer. It also turned out that when they have the wrong perspective, they take more time than others, repeating the same operation over and over again. A model proposed in the paper succeeds in estimating understanding using the perspective status estimated by the hidden Markov model and its frequency as explanatory variables. The results of the cross-validation showed that the model almost perfectly predicted the understanding labels given by the instructors. Therefore, this tool can provide appropriate guidance to those who lack a clear perspective or those who have a clear perspective but are repeating the same operation. The previous method we proposed for analyzing understanding using code puzzles can inform the instructor of the characteristics of learners, such as those who do not read the problem text or those whose drawings are too guesswork. We can expect to get additional feedback by applying it at the same time as the method proposed in this paper. To accurately measure programming ability, it would be necessary to be able to extract the answering process and its behavior in an environment where one can focus on combining programming elements like a code puzzle. Applying this method to a multi-participant lecture, it is possible to easily grasp the status of the learners' perspective and the key points they do not understand to automatically find the learners who need help by estimating their level of understanding.

References

1. Hellas A, Ihantola P, Petersen A, Ajanovski VV, Gutica M, Hynninen T, Knutas A, Leinonen J, Messom C, Liao SN (2018) Predicting academic performance: a systematic literature review. In: Proceedings companion of the 23rd annual ACM conference on innovation and technology in computer science education, ITiCSE 2018 companion. Association for Computing Machinery, New York, NY, USA, pp 175–199. https://doi.org/10.1145/3293881.3295783
2. Nikula U, Gotel O, Kasurinen J (2011) A motivation guided holistic rehabilitation of the first programming course. ACM Trans Comput Educ 11(4). https://doi.org/10.1145/2048931.2048935
3. Corney M, Teague D, Thomas RN (2010) Engaging students in programming. In: Proceedings of the twelfth Australasian conference on computing education, vol 103, ACE'10. Australian Computer Society, Inc., AUS, pp 63–72
4. Whipkey KL (1984) Identifying predictors of programming skill. SIGCSE Bull 16(4):36–42. https://doi.org/10.1145/382200.382544
5. Mazlack LJ (1980) Identifying potential to acquire programming skill. Commun ACM 23(1):14–17. https://doi.org/10.1145/358808.358811
6. Nesbit JC, Liu QLA, Adesope OO (2015) Work in progress: intelligent tutoring systems in computer science and software engineering education. In: Proceeding 122nd American society for engineering education annual conference

7. Luo J, Wang T (2020) Analyzing students' behavior in blended learning environment for programming education. In: Proceedings of the 2020 the 2nd World symposium on software engineering, WSSE 2020. Association for Computing Machinery, New York, NY, USA, pp 179–185. https://doi.org/10.1145/3425329.3425346
8. Schnotz W, Kürschner C (2007) A reconsideration of cognitive load theory. Educ Psychol Rev 19(4):469–508. https://doi.org/10.1007/s10648-007-9053-4
9. Parsons D, Haden P (2006) Parson's programming puzzles: a fun and effective learning tool for first programming courses. In: Proceedings of the 8th Australasian conference on computing education, vol 52, ACE'06. Australian Computer Society, Inc., AUS, pp 157–163
10. Crow T, Luxton-Reilly A, Wuensche B (2018) Intelligent tutoring systems for programming education: a systematic review. In: Proceedings of the 20th Australasian computing education conference, ACE'18. Association for Computing Machinery, New York, NY, USA, pp 53–62. URL https://doi.org/10.1145/3160489.3160492
11. Jadud MC (2006) Methods and tools for exploring novice compilation behaviour. In: Proceedings of the second international workshop on computing education research, ICER'06. Association for Computing Machinery, New York, NY, USA, pp 73–84. https://doi.org/10.1145/1151588.1151600
12. Mysore A, Guo PJ (2018) Porta: profiling software tutorials using operating-system-wide activity tracing. In: Proceedings of the 31st annual ACM symposium on user interface software and technology, UIST'18. Association for Computing Machinery, New York, NY, USA, pp 201–212. https://doi.org/10.1145/3242587.3242633
13. Guo PJ (2015) Codeopticon: real-time, one-to-many human tutoring for computer programming. In: Proceedings of the 28th annual ACM symposium on user interface software & technology, UIST'15. Association for Computing Machinery, New York, NY, USA, pp 599–608. https://doi.org/10.1145/2807442.2807469
14. Asai S, Phuong DTD, Shimakawa H (2019) Identification of factors affecting cognitive load in programming learning with decision tree, pp 624–633. https://doi.org/10.17706/jcp.14.11.624-633
15. Ihantola P, Sorva J, Vihavainen A (2014) Automatically detectable indicators of programming assignment difficulty. In: Proceedings of the 15th annual conference on information technology education, SIGITE'14. Association for Computing Machinery, New York, NY, USA, pp 33–38. https://doi.org/10.1145/2656450.2656476
16. Villamor M (2020) A review on process-oriented approaches for analyzing novice solutions to programming problems. Res Pract Technol Enhanced Learn 15. https://doi.org/10.1186/s41039-020-00130-y
17. Watson C, Li FW, Godwin JL (2013) Predicting performance in an introductory programming course by logging and analyzing student programming behavior. In: 2013 IEEE 13th International conference on advanced learning technologies, pp 319–323. https://doi.org/10.1109/ICALT.2013.99
18. Watson C, Li FW, Godwin JL (2014) No tests required: comparing traditional and dynamic predictors of programming success. In: Proceedings of the 45th ACM technical symposium on computer science education, SIGCSE'14. Association for Computing Machinery, New York, NY, USA, pp 469–474. https://doi.org/10.1145/2538862.2538930
19. Tabanao ES, Rodrigo MMT, Jadud MC (2011) Predicting at-risk novice java programmers through the analysis of online protocols. In: Proceedings of the seventh international workshop on computing education research, ICER'11. Association for Computing Machinery, New York, NY, USA, pp 85–92. https://doi.org/10.1145/2016911.2016930
20. Futamura Y, Kawai T, Horikoshi H, Tsutsumi M (1981) Development of computer programs by problem analysis diagram (pad). In: Proceedings of the 5th international conference on software engineering, ICSE'81. IEEE Press, pp 325–332. https://doi.org/10.5555/800078.802545
21. Blikstein P, Worsley M, Piech C, Sahami M, Cooper S, Koller D (2014) Programming pluralism: using learning analytics to detect patterns in the learning of computer programming. J Learn Sci 23(4):561–599. https://doi.org/10.1080/10508406.2014.954750

Camera and LiDAR Fusion for Point Cloud Semantic Segmentation

Ali Abdelkader and Mohamed Moustafa

Abstract Perception is a fundamental component of any autonomous driving system. Semantic segmentation is the perception task of assigning semantic class labels to sensor inputs. While autonomous driving systems are currently equipped with a suite of sensors, much focus in the literature has been on semantic segmentation of camera images only. Research in the fusion of different sensor modalities for semantic segmentation has not been investigated as much. Deep learning models based on transformer architectures have proven successful in many tasks in computer vision and natural language processing. This work explores the use of deep learning transformers to fuse information from LiDAR and camera sensors to improve the segmentation of LiDAR point clouds. It also addresses the question of which fusion level in this deep learning framework provides better performance.

Keywords Point cloud semantic segmentation · Sensor fusion · Vision transformers · Computer vision

1 Introduction

Autonomous vehicles have input data that come from a suite of sensors mounted on the vehicle such as radar, camera, and LiDAR. In recent years, deep learning models have achieved state of the art in semantic segmentation in 2D image data from the camera. However, semantic segmentation of 3D data such as LiDAR point cloud has proven more challenging. Point cloud semantic segmentation is the task of assigning class label to every point in the input point cloud. Fusing multiple sensor data for point cloud segmentation is a lesser investigated topic than 2D or 3D segmentation. This work aims at developing deep learning models that would take a LiDAR point cloud along with its associated RGB image and perform semantic segmentation

A. Abdelkader (✉) · M. Moustafa
American University in Cairo, Cairo, Egypt
e-mail: aly.abdelkader@aucegypt.edu

M. Moustafa
e-mail: m.moustafa@aucegypt.edu

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_45

of the point cloud. More specifically, the deep learning models used in this work are based on the popular transformers architecture. Transformers models are popular because of their computational efficiency and scalability which allow for the training of massive models of over 1B parameters [2]. They have proven very successful in natural language processing tasks. They have been recently introduced to computer vision tasks by Dosovitskiy et al. [2]. There are three broad techniques for fusion of sensor modalities in deep learning: early, middle, and late fusion. Early fusion combines data from camera and LiDAR at the input of the model. Middle fusion extracts features from each modality separately and combines them in feature space before further processing. Late fusion applies different models for each modality separately, then, learns to combine decisions for final labels predictions. This work aims at answering the following questions:

1. *Do fusion models benefit from features extracted by image-based transformer for point cloud semantic segmentation ?*
2. *Which level of fusion between image-based transformer and LiDAR segmentation network performs best ?*

2 Related Work

Transformer Models Transformer models have been introduced in the seminal paper 'Attention Is All You Need' [14]. They have proven very successful in natural language processing tasks. Dosovitskiy et al. [2] ported the original transformer model of [14] to image classification task calling it vision transformer (ViT). They concluded that the inductive bias baked into the structure of regular CNN networks allows them to perform favorably when training is done using small datasets. However, the story changes as the size of training dataset increases because the large-scale training of ViT outperforms inductive bias CNNs of similar network size [14]. Touvron et al. [13] addressed the large-scale training requirement of ViT. They showed that vision transformers can achieve competitive performance using training on ImageNet only by using advanced regularization and data augmentation techniques. Touvron et al. [13] also introduced a distillation strategy that would allow for learning to transfer from a pretrained teacher CNN model to vision transformer during training. In this strategy, the regular cross-entropy is augmented with cross-entropy loss calculated between the embedding of a newly introduced distillation token and the predictions of teacher model.

Point Cloud Segmentation Point cloud segmentation models can be categorized by how they present the point cloud to a deep learning model. Point-based representation approaches deal with unordered point clouds directly. The most popular model in this category is PointNet[9]. The basic idea behind PointNet [9] is to learn an order invariant representation of point cloud using shared multilayer perceptrons (MLPs) and symmetric functions such as max pooling layers. Point-based representations approaches tend to perform well on sparse point clouds; however, they

tend to suffer computationally as input point clouds become denser [1]. On the other hand, volumetric approaches such as SEGCloud [12], SPLATNet [10], discretize the 3D space of the point clouds into smaller units such as voxels or lattice. 3D convolutional networks are then applied on this representation to produce the final semantic label. Those approaches usually come with heavy computational cost due to their 3D operations. Liu et al. [7] found out that the irregularity of the LiDAR point cloud forces point-based models to waste up to 80% of their computations performing expensive neighborhood clustering and to exhibit cache inefficiency as points are not stored contiguously in memory [7]. On the other hand, voxel-based models have a favorable memory locality characterizes, but, suffer from prohibitively large memory footprint which forces those models to perform aggressive downsampling and operate at a low resolution [7]. As a result, Liu et al. [7] introduced the point voxel convolution primitive. Point voxel convolutions take the best of both worlds by processing point clouds using two branches: a voxel-based branch and a point-based branch. The voxel-based branch is designed to capture coarse features by performing 3D convolutions and decovolutions on the voxel representation of the point cloud. The point-based branch is designed to capture fine-grained features by operating on each point individually. However, representing LiDAR point cloud in a 3D voxel space as regular tensors is memory inefficient as many of those voxels end up empty, not storing valid data. Tang et al. [11] introduced spare point voxel convolution by replacing regular 3D convolution with sparse convolutions which convolve filters' kernels with the input if the input has valid data, thus ignoring empty entries.

Fusion Models Compared to 2D and 3D segmentation models, fusion models are far less explored in the literature. Fusion models can be categorized based on their fusion technique. There are three broad techniques to fuse sensor modalities: early, middle, and late fusion. An example of early fusion model is superpoint graph [6]. Its input is the set of points in a point cloud and their associated RGB values. Superpoint proceeds first by partitioning the point cloud into neighboring points forming superpoints. Then, it creates a superpoint graph by forming annotated edges between superpoints. It uses a PointNet network and a gated graph network to produce its final class predictions. Meyer et al. [8] proposed a middle fusion model in which image features were extracted by a regular ResNet and added to the LiDAR polar grid map using the relative calibration matrix between the two sensors. Krispel et al. [5] took this step further by proposing to wrap, not only the final feature vector from the image encoder network but also features from multiple layers of the encoder in to the LiDAR segmentation network using first-order polyharmonic spline interpolation to establish a correspondence between the LiDAR and the image feature maps' coordinates. An example of late fusion is xMUDA [4] which is targeting the domain adaptation task. In their model, the RGB image and LiDAR point cloud go through independent multi-headed networks. Fusion is primarily done in the loss objective. Regular cross-entropy loss is calculated using predictions from the first head of given modality and ground truth label [4]. Kullback–Leibler divergence loss is calculated using predictions from the second head of that modality and the predictions from the first head of the other modality [4].

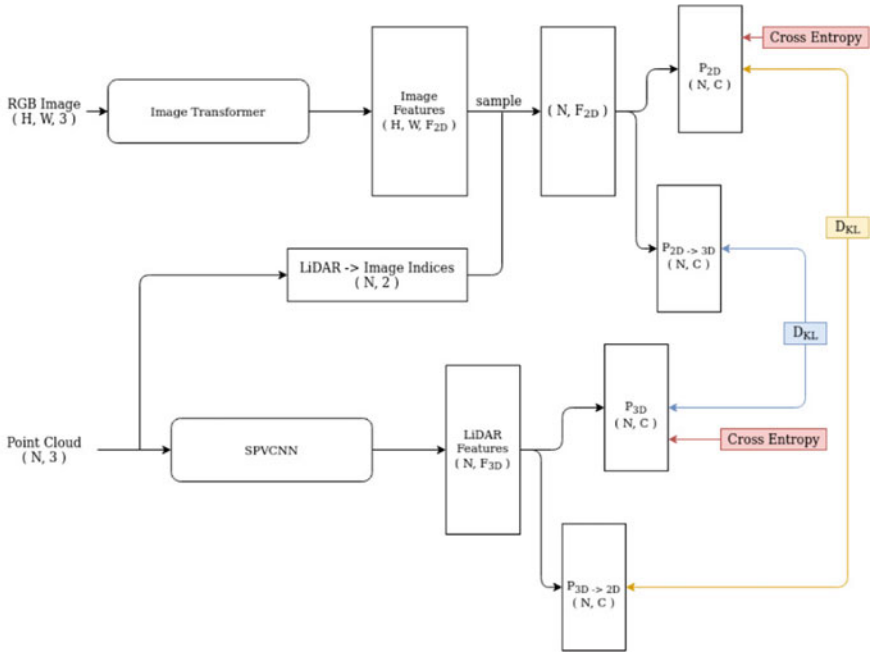


Fig. 1 General training framework, similar to xMUDA [4]

3 Proposed Approach

This work aims at investigating the use of image-based transformers in fusion frameworks for point cloud semantic segmentation. It adopts similar framework to xMUDA [4] which has two branches for image and LiDAR modalities. In this section, the different components of this framework will be explained starting with the input representation (Fig. 1).

3.1 Input Representation

LiDAR point cloud is represented as a sparse voxelized tensor. The transformation of the regular point cloud into sparse voxelized tensor is adopted from [11]. The 2D pixel coordinate for each 3D LiDAR point can obtained by $x = PX$ where x is the pixel coordinate, P is the 3×4 projection matrix provided by SemanticKITTI, and X is the LiDAR point 3D coordinates in homogeneous space. Those LiDAR $\rightarrow$ image indices are stored in order to be used later to sample features generated from the image transformer. The RGB image is represented as a matrix of shape (H, W, 3) where W and H are the image width and height respectively.

3.2 Models

There are two backbone models for processing each modality: camera and LiDAR. They will also be used to create the three fusion models for each fusion level.

Image Transformer The vision transformer (ViT) is the main powerhouse of image-based branch. It is responsible for extracting features from images that would aid in the semantic segmentation of the LiDAR point cloud. The vision transformer used in this work is the distilled data-efficient vision transformer (DeiT) from [13]. The image provided to the transformer is divided into patches. The number of patches is determined by the patch resolution (P, P) resulting in $N_{\text{patches}} = \text{HW}/P^2$. Each patch of shape $(P^2, 3)$ is then linearly embedded into a vector of shape (P^2, F_{2D}). Positional embeddings are added to these patches in order to provide the transformer model with the relative spatial information of the patches that was lost when the image was divided. In addition to the image patches, the DeiT model takes two additional tokens: class token, distillation token. Those tokens were used during the pretraining of the DeiT model on ImageNet. Since we are interested only in features extracted from the image patches, those additional tokens are ignored. The DeiT transformer has 11 transformer blocks stacked on top of each other. Multi-head self-attention (MSA) is the core component for the transformer block [13]. Its working procedure is as follows:

1. The input is multiplied by weight matrices W^Q,W^K,W^V forming query Q, keys K, and value V matrices, respectively
2. Applying the softmax function to the normalized inner product of query and keys produces the weight scores.
3. Attention head's output is the value matrix weighted by the scores [13]

$$Attention(\mathsf{Q}, \mathsf{K}, \mathsf{V}) = Softmax(\frac{\mathsf{Q}\mathsf{K}^\mathsf{T}}{\sqrt{d}})\mathsf{V}. \tag{1}$$

4. Steps 1 to 3 are repeated for each attention head resulting in an output matrix of shape (N_{patches}, hd) where h is the number of attention heads and d is the output size of each head. This output matrix is linearly projected to shape $(N_{\text{patches}}, F_{2D})$ [13].

Multi-layer perceptrons' (MLP) network which has two neural layers with GELU activation is added on top of the MSA. The first layer expands the embedding dimension from F_{2D} to $4 \times F_{2D}$. The second squeezes it back to F_{2D} [13]. Skip connections and layer normalization are added to turn MSA and FFN into residual operators [13].

LiDAR backbone This work adopts sparse voxel convolutional neural network (SPVCNN) from [11] as LiDAR backbone model. SPVCNN uses sparse point voxel convolution as it is primarily primitive. Point voxel convolution uses two branches to process the 3D input. A point-based branch is a multilayer perceptron network that acts on each point in the point cloud individually. This branch helps the model maintain high-resolution representation for each point [11]. This improves the model's

performance on classes which are usually represented with very few points such as pedestrians and cyclists [11]. The second branch is the voxel branch. The voxel branch converts the point cloud into a 3D voxel representations. It uses 3D convolutions to capture global features of input scene. SPVCNN [11] replace regular 3D convolutions with their sparse counterparts. Sparse convolutions create a kernel map between the convolution's input and output. They operate only on outputs that belong to a valid input, thereby, dramatically improving memory efficiency because they ignore input voxels that are empty. The LiDAR point cloud is feed to the model as a sparse tensor as discussed earlier. On the voxel branch, the 3D encoder applies residual 3D sparse convolutions to the input [11]. The output of the encoder is devoxelized using trilinear interpolation and fused with the output of MLP on the point-based branch [11]. Similarly, the decoding steps apply 3D sparse deconvolutions to upsample back the voxel representation while fusing with the point-based branch [11].

Fusion Models The image transformer and LiDAR backbone model were used as building blocks to create separate fusion models representing the different fusion levels. **Early Fusion** fuses features from both modalities at the early stages of the pipeline. For our early fusion model, features from the first transformer block in the image backbone are sampled using LiDAR $\rightarrow$ image indices resulting in a tensor shape (N, F_{2d}) containing features for each LiDAR point. This tensor is fused into the LiDAR backbone model by adding it to the output of the stem block of the latter. On the other hand, **Middle Fusion** fuses features from both modalities at a later stage of the deep learning pipeline. For our middle fusion model, features from the 6^{th} transformer block are sampled in a similar way to the early fusion model. The resulting tensor is fused into the LiDAR backbone model by adding it to the output of the 3D encoder (before the deconvolutional layers). **Late Fusion** fuses features from both modalities after each modality's pipeline has produced its semantic predictions independently. In this work, late fusion is done at the loss stage. As will be discussed later, the Kullback–Leibler loss terms $D_{\mathrm{KL}}(\mathsf{P}_{\mathrm{3D}}||\mathsf{P}_{\mathrm{2D}\rightarrow\mathrm{3D}})$ and $D_{\mathrm{KL}}(\mathsf{P}_{\mathrm{2D}}||\mathsf{P}_{\mathrm{3D}\rightarrow\mathrm{2D}})$ are responsible for performing the late fusion between the different modalities.

3.3 Outputs

The LiDAR backbone outputs a features map of shape (N, F_{3D}) where F_{3D} is the number of feature channels of the model. This features map goes through two linear heads. The first head generates tensor P_{3D} of shape (N, C) representing the class scores for each point in N. The second head generates another tensor $\mathsf{P}_{\mathrm{3D}\rightarrow\mathrm{2D}}$ of the same shape, but it represents the class that the 3D network thinks the 2D network would predict for each point. On the other hand, the camera image is provided to the image transformer which generates feature map of shape (H, W, F_{2D}) where F_{2D} is number of output features channels of transformer. This feature map is sampled using the LiDAR $\rightarrow$ image indices which results in a feature vector of shape (N, F_{2D}) where N is the number of points in point cloud. Similarly, this feature vector goes through two linear heads generating P_{2D}, $\mathsf{P}_{\mathrm{2D}\rightarrow\mathrm{3D}}$.

3.4 Objective Function

Each sensor modality optimizes a separate but analogues objective. The overall objective function for LiDAR modality is given by

$$\min_{\theta}(1-\alpha) \times CrossEntropy(\mathsf{P}_{3D}, \mathbf{G}_t) + \alpha \times D_{KL}(\mathsf{P}_{2D}||\mathsf{P}_{3D\to 2D}) \tag{2}$$

$CrossEntropy(\mathsf{P}_{3D}, \mathbf{G}_t)$ is a cross-entropy loss between the LiDAR model predictions and the ground truth labels for each LiDAR point. In our experiments, similar to xMUDA [4], the cross-entropy loss terms were also weighted using log-smoothed class weights.

$$CrossEntropy(\mathsf{P}_{3D}, \mathbf{G}_t) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} \mathbf{G}_t^{(i,c)} \log \mathsf{P}_{3D}^{(i,c)} \tag{3}$$

$D_{KL}(\mathsf{P}_{2D}||\mathsf{P}_{3D\to 2D})$ is the Kullback–Leibler divergence between the actual predictions of image transformer and the LiDAR model's predictions for it. This is given by

$$D_{KL}(\mathsf{P}_{2D}||\mathsf{P}_{3D\to 2D}) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} \mathsf{P}_{2D}^{(i,c)} \log \frac{\mathsf{P}_{2D}^{(i,c)}}{\mathsf{P}_{3D\to 2D}^{(i,c)}} \tag{4}$$

4 Experiment

Dataset SemanticKITTI is based on the odometry KITTI dataset. It provides 22 sequences (11 for training and 11 for testing) with high quality point-wise annotations. It has 19 different class labels. However, the labels for the testing sequences are not available for public. They are only reserved for submission on the dataset's benchmark. In our experiments, the training sequences 0–11 were used. Performance was evaluated by mean Intersection over Union (mIoU) on validation sets that were created following 3-fold cross-validation strategy.

Data Preprocessing SemanticKitti provides LiDAR point cloud for each scene in which each point has cartesian coordinates x, y, z and $intensity$. The pixel coordinates for each point are calculated using the 3×4 projection matrix provided by the dataset's authors. As part of prepossessing steps, points that lie outside the camera frame are filtered out. SemanticKitti also provides RGB images of shape (1226, 370, 3). However, the pretrained DeiT expects images in shape (384, 384, 3). Therefore, the input images were resized using bilinear interpolation to the size expected by the image transformer and resized back afterward.

Training All models were built using PyTorch framework. The LiDAR backbone is based on the original implementation of SPVCNN provided by the authors of [11]. The image transformer is based on an implementation of DeiT model from

the popular Timm library [15]. The Timm library also provided the ImageNet pretrained weights for this model. Models were trained on the training sets with Adam optimizer on Nvidia DGX server with alpha α, patch Size (P), learning rate, L2 regularization and number of epochs set to 0.1, 16, 1e-4, 0.005, and 50 respectively. For the KL divergence terms in the objective functions, the target variable was detached in PyTorch to backpropagate in either the image or the LiDAR network [4]. Similarly, in the fusion models, features from the image modality were detached in PyTorch after the sampling step to only backpropagate in LiDAR network when computing gradients for the 3D loss.

5 Results and Discussion

The results in Table 1 were obtained by averaging the performance of each model across the validation sets following the 3-fold cross-validation strategy. The results show that the fusion models outperformed the baseline LiDAR-only model on mIoU by a slight margin. The reason for the improvement being only minor can be attributed to the fact that the DeiT was pretrained on ImageNet which is approx 1M samples while SemanticKitti is approx 20k samples. Therefore, it could be due to the DeiT overfitting to the small training sets (relative to ImageNet training) in our experiments. Additionally, in our experiments the fusion models outperformed the baseline in all of the 19 classes. This is an indication that fusion models can benefit from features extracted by an image transformer. Moreover, given that DeiT is known to scale well with dataset size, we predict that the performance between LiDAR and fusion models would widen with dataset size. Unfortunately, all models failed to motorcyclist class. This can attributed to the class imbalance in SemanticKitti dataset in which motorcyclist class is one of the least represented classes.

6 Conclusion and Future Work

Future Work Although our experiments show promise for fusion models, there are several ways to improve them. First, training with a larger dataset would help better discern the benefits of fusion over using only one modality. Second, in our experiments, we have chosen a specific architecture to perform middle fusion. However, as mentioned in [3], there are multiple ways to design middle fusion architectures. Future experiments can also explore such different architectures.

Conclusion This work explored using deep learning to combine information from LiDAR and camera sensors to improve the segmentation of LiDAR point clouds. More specifically, it aimed at investigating the role that the recent advances of image-based transformer architectures can play in a fusion framework with LiDAR-based models. Our experiments show slight improvement in performance of such fusion models over LiDAR-only baseline. This points at the potential that image-based

Table 1 SemanticKitti dataset results (19 classes) for all models on sequence 08 (test set)

Class	LiDAR	Early fusion	Middle fusion	Late fusion
car	0.907	0.909	**0.915**	0.913
bicycle	0.074	**0.139**	0.112	0.096
motorcycle	0.538	0.504	**0.548**	0.532
truck	0.184	0.185	**0.248**	0.202
other-vehicle	0.290	0.262	**0.295**	0.281
person	0.409	0.455	**0.493**	0.375
bicyclist	0.409	0.377	0.457	**0.473**
motorcyclist	0.000	0.000	0.000	0.000
road	0.939	**0.944**	0.942	0.941
parking	0.448	**0.475**	0.474	0.448
sidewalk	0.759	**0.775**	0.768	0.761
other-ground	0.007	0.005	**0.008**	0.004
building	0.852	**0.863**	0.860	0.854
fence	0.609	0.613	**0.616**	0.613
vegetation	0.827	**0.832**	0.826	0.827
trunk	0.641	0.633	0.637	**0.647**
terrain	0.658	**0.668**	0.653	0.652
pole	0.582	0.588	**0.587**	0.596
traffic-sign	0.513	0.514	**0.518**	0.509
mIoU	0.482	0.487	**0.498**	0.486

transformers can play in fusion models in the future. Our experiments also show that middle-level fusion outperforms the other types of fusion models which hints that fusion models learn more from features extracted from the image as opposed to final semantic prediction from an independent image-based pipeline.

References

1. Behley J, Garbade M, Milioto A, Quenzel J, Behnke S, Stachniss C, Gall J (2019) SemanticKITTI: a dataset for semantic scene understanding of LiDAR sequences. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)
2. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N (2021) An image is worth 16x16 words: transformers for image recognition at scale. In: International conference on learning representations
3. Feng D, Haase-Schütz C, Rosenbaum L, Hertlein H, Glaeser C, Timm F, Wiesbeck W, Dietmayer K (2020) Deep multi-modal object detection and semantic segmentation for autonomous driving: datasets, methods, and challenges. IEEE Trans Intelligent Transp Syst 22(3):1341–1360

4. Jaritz M, Vu TH, Charette R, Wirbel E, Perez P (2020) xmuda: cross-modal unsupervised domain adaptation for 3d semantic segmentation. In: Proceedings of the IEEE/CVF vonference on Computer Vision and Pattern Recognition (CVPR)
5. Krispel G, Opitz M, Waltner G, Possegger H, Bischof H (2020) Fuseseg: lidar point cloud segmentation fusing multi-modal data. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, pp 1874–1883
6. Landrieu L, Simonovsky M (2018) Large-scale point cloud semantic segmentation with superpoint graphs. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 4558–4567 https://doi.org/10.1109/CVPR.2018.00479
7. Liu Z, Tang H, Lin Y, Han S (2019) Point-voxel cnn for efficient 3d deep learning. In: Advances in neural information processing systems
8. Meyer GP, Charland J, Hegde D, Laddha A, Vallespi-Gonzalez C (2019) Sensor fusion for joint 3d object detection and semantic segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops
9. Qi CR, Su H, Mo K, Guibas LJ (2017) Pointnet: deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 652–660
10. Su H, Jampani V, Sun D, Maji S, Kalogerakis E, Yang MH, Kautz J (2018) Splatnet: sparse lattice networks for point cloud processing. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2530–2539
11. Tang H, Liu Z, Zhao S, Lin Y, Lin J, Wang H, Han S (2020) Searching efficient 3d architectures with sparse point-voxel convolution. In: European conference on computer vision
12. Tchapmi LP, Choy CB, Armeni I, Gwak J, Savarese S (2017) Segcloud: semantic segmentation of 3d point clouds. In: International conference on 3D Vision (3DV)
13. Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jégou H (2021) Training data-efficient image transformers & distillation through attention. In: International conference on machine learning, PMLR, pp 10,347–10,357
14. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. In: Advances in neural information processing systems, pp 5998–6008
15. Wightman R (2019) Pytorch image models. https://github.com/rwightman/pytorch-image-models. https://doi.org/10.5281/zenodo.4414861

A Priori Study on Factors Affecting MapReduce Performance in Cloud-Based Environment

Vandana Vijay and Ruchi Nanda

Abstract In the current era, global data have been rising at a very fast speed due to the excessive use of technologies including cloud and IoT. It leads to the development of big data that can handle and analyze a high volume of data regularly. Cloud computing provides a reliable, available, and scalable environment for the processing of this huge data. MapReduce has become an important computing model for processing and generating high-volume datasets on a cluster of machines. It not only allows distributed processing but also significant attributes like flexibility, versatility, load adjusting, and adaptation to internal failure. Despite these benefits, the performance of this framework gets affected by multiple factors, namely indexing, data skew, joining, caching, and load balancing. The main objective of this paper is to identify the factors to resolve its performance issues and investigate alternate strategies to improve the MapReduce query performance in the cloud-based environment.

Keywords Big data · MapReduce · Hadoop · HDFS · Cloud computing

1 Introduction

Excessive use of social networking sites, smartphones, laptops, computers, and other handheld devices leads to the accumulation of a huge volume of data on the Internet. The main demand is to access this data in an efficient manner. MapReduce has raised up as a worldwide tool for the rapid execution of huge textual data, images, and videos. Since MapReduce's current design does not consider virtualization, so cloud computing provides a virtualized environment to it. Cloud computing provides the facility of automatic provisioning of computing resources as per the demand of the users. Hence, running MapReduce on cloud computing is becoming popular [3] as it supplies a reliable, available, and scalable environment for the processing of huge data. But, there are various other factors that affect the performance of MapReduce in a cloud-based environment. The main aim of this paper is to highlight those factors

V. Vijay · R. Nanda (✉)
Department of CS and IT, IIS, Deemed to Be University, Jaipur, India
e-mail: ruchi.nanda@iiusniv.ac.in

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_46

which facilitate enhancing the query performance of MapReduce in the cloud environment. The paper is divided into four major sections: The first section describes the concepts of MapReduce and its open-source implementation, i.e., Apache Hadoop. The second section elaborates on the factors affecting MapReduce's performance in the cloud-based environment. The third section provides a comprehensive review of the literature. The last section describes the conclusions drawn from the literature review and summarizes the alternate strategies adopted by various researchers to enhance its performance. This section also highlights the areas where further research can be taken up in the future.

1.1 MapReduce

MapReduce is a programming-type framework. It is designed for the execution of parallel and distributed high-volume data using multiple nodes. The open-source implementation of MapReduce is Apache Hadoop, which is mainly used for distributed storage and processing of data. Hadoop distributed file system (HDFS) is used for storage while MapReduce for executing the data. Hadoop follows the concept of master and slave as shown in Fig. 1. Masters include NameNode, SecondaryNode, and JobTracker, while slave includes DataNode and TaskTracker. HDFS deals with NameNode, SecondaryNode, and DataNode, while MapReduce

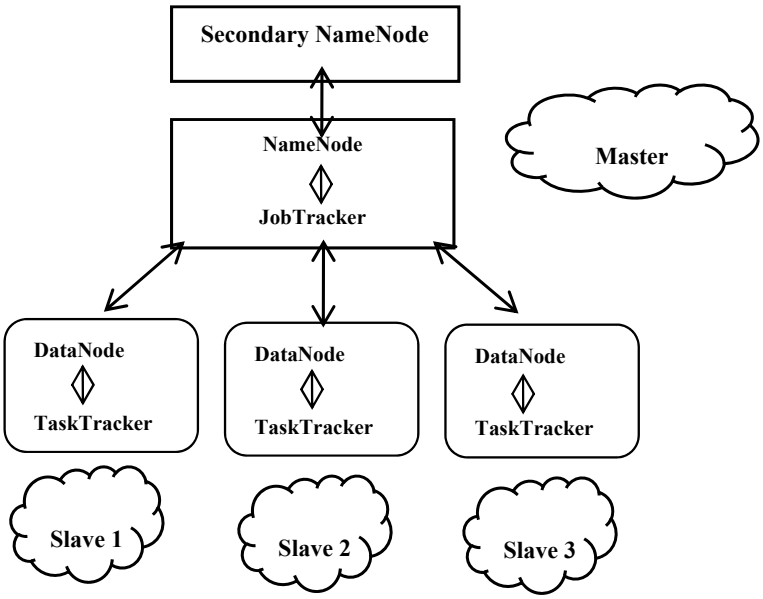


Fig. 1 Hadoop master/slave structure

deals with JobTracker and TaskTracker. MapReduce processes the big data in the given manner.

Firstly, the job requests are received by the JobTracker. It, later on, splits the input data and assigns the tasks to the TaskTracker for processing. Secondly, the JobTracker monitors the progress of each task and also handles any type of failure. The TaskTracker always performs the tasks under the supervision of the JobTracker. These tasks are executed through a Map or Reduce function [6]. During the processing of tasks, there are multiple factors that affect the performance of the MapReduce framework directly or indirectly. These factors are explained in the next section.

2 Factors Affecting Query Performance of MapReduce

Different factors may affect the performance of MapReduce. Among them, some of the factors may reside in the MapReduce architecture, while the others may be external factors.

The factors that may have a significant effect on the observed performance of MapReduce are given as:

(a) Indexing is used to maximize the database performance. It reduces thedisk access required during query processing through MapReduce.
(b) Data skew is the imbalance in allocating the data for processing to each MapReduce task.
(c) Joining facilitates the retrieval of data from two different relations based on a Cartesian product of the two relations.
(d) Caching represents the intermediate data generated by MapReduce tasks during processing.
(e) Input data parsing is a conversion-type process. While retrieving data, it converts raw data into the key/value pairs.
(f) Data size is directly related to the runtime of a job. It includes input, output, and shuffle data.
(g) Load balancing is a technique that divided the work equally between all the mappers and reducers units.
(h) Data availability along with high-speed access must be ensured from the underlying storage system during data retrieval by MapReduce (Input data storage).
(i) Input/output mode is used to retrieve the data from any storage system. Important modes for reading the data are (1) Direct I/O which reads the data from the local cache to memory, (2) Streaming I/O which reads the data through inter-process communication from storage.

3 Review of Literature

This section refers to the crux of the research papers published in the field of MapReduce concerning cloud computing. An attempt is made to throw light on the factors affecting the performance of MapReduce in a cloud-based environment. It has been presented in Table 1.

4 Conclusion and Future Scope

In this paper, several factors have been investigated which affect MapReduce performance significantly and also investigate alternate implementation schemes for each factor proposed by the various researchers to enhance its performance. Researchers have developed various solutions like B-tree index, HAIL—Hadoop aggressive indexing library, NPIY—novel partitioner, indexed MapReduce join algorithm, binary theta-join, multi-way theta-join, autocache, and task failure resilience (TFR) algorithm. Through this paper, an attempt is made to identify all the factors which affect MapReduce performance in a theoretical manner. As MapReduce is at its peak, so there is a scope in the coming future to further improve its performance by altering variables such as memory settings, job or task scheduling, data locality, and system latency. These directions will serve as a prospective road map for upcoming research.

Table 1 Review of literature

Papers authors	Proposed/developed methodology	Factors identified	Conclusion drawn
Kavitha and Anita [6]	Proposed task failure resilience (TFR) algorithm to implement fault tolerance in MapReduce	Caching	Enhances the Hadoop performance through the fast healing of the intermediate results from the in-memory data store
Mittal et al. [11]	Proposed B-tree index for the input dataset. Developed a parallel B-tree index from the traditional B-tree index	Indexing	The efficiency of random processing of data in Hadoop can be improved by making an index on the input dataset
Dittrich et al. [2]	Proposed Hadoop aggressive indexing library (HAIL), to create clustered indexes data block	Indexing	HAIL improves up to 60% over HDFS. It executes 68 × faster than Hadoop
Irandoost et al. [4]	Proposed learning automata hash partitioner (LAHP) algorithm for custom distribution of intermediate key-value results to reducers	Data skew	LAHP distributes the load more equitably to the reducers with an accuracy of 99%
Lu et al. [9]	Proposed novel partitioner for improving MR performance (NPIY). It provides an even distribution of intermediate data with negligible overhead	Data skew	Performance improves 41.66%—homogeneous 58.68%-heterogeneous clusters
Liu and Li [10]	Proposed two methods, namely max–min value-based filter method (MMF) and divide and merge (DM) method	Joining	It speeds up the query efficiency from 1.6 up to 77.4. But, this efficiency can vary with the theta-join situation
Khafagy [7]	Proposed indexed MapReduce join algorithm. It used index in the large table to decrease I/O and shuffling	Joining	This algorithm has higher performance than others when increasing the data size from 100 to 500 million records
Herodotou [3]	Proposed autocache for automated cache management in distributed file system	Caching	It increases the system performance by automatically deciding when and which file to store and evict from the cache
Jiang et al. [5]	They identified factors, data parsing, and indexing. Investigate alternative implementation strategies for each factor	Input data parsing	The results show that MapReduce performance improved by a factor of 2.5 to 3.5

(continued)

Table 1 (continued)

Papers authors	Proposed/developed methodology	Factors identified	Conclusion drawn
Tang and Bressoud [13]	They identified that the input size depends on the total job time of a MapReduce task	Data size	To reduce the overheads, configuration parameters are adjusted. For instance, block size or level of concurrency. In this way, one can utilize the Hadoop cluster fully
Patil et al. [12]	Secured Hadoop in the form of cloud service is proposed that achieves load balancing and quick process of huge data in less amount of time	Load balancing	It improves security problems and achieves load balancing and a quick process of huge data in less amount of time
Basha et al. [1]	Concise the storage and execution speed in the enhanced cloud along with the Hadoop	Input data storage	Sqoop loads data from DB to HDFS while Flume from server to HDFS
Kumar and Gupta [8]	Proposed algorithm for assigning more read/write requests to higher I/O speeds data nodes and assigning fewer read/write requests to lower I/O speeds data nodes	I/O mode	It showed an improvement of 23.27% as compared to the standard HDFS and 14.62% from the HDFS balancer

References

1. Basha S, Kumar P, Babu S (2016) Storage and processing speed for knowledge from enhanced cloud computing with hadoop framework: a survey. IJSRSET 2(2):126–132
2. Dittrich J, Quiané-Ruiz J, Richter S, Schuh S, Jindal A, Schad J (2012) Only aggressive elephants are fast elephants. Proc VLDB Endow 5:1591–1602
3. Herodotou H (2019) AutoCache: employing machine learning to automate caching in distributed file systems. In: 35th international conference on data engineering workshops. IEEE, pp 133–139
4. Irandoost M, Rahmani A, Setayeshi S (2019) A novel algorithm for handling reducer side dataskew in MapReduce based on a learning automata game. Inf Sci 501:662–679
5. Jiang D, Ooi B, Shi L, Wu S (2010) The performance of MapReduce: an in-depth study. Proc VLDB Endow 3(1):472–483
6. Kavitha C, Anita X (2020) Task failure resilience technique for improving the performance of MapReduce in hadoop. ETRI J 42(5):748–760
7. Khafagy M (2015) Indexed map-reduce join algorithm. Int J Sci Eng Res 6(5):705–711
8. Kumar V, Gupta H (2017) Performance improvement of mapreduce through enhancement in hadoop block placement algorithm. Int J Comput Sci Mob Comput 6(4):389–396
9. Lu H, Hai-Shan C, Ting-Ting H (2012) Research on hadoop cloud computing model and its applications. In: 2012 third international conference on networking and distributed computing. IEEE, pp 59–63
10. Liu W, Li Z (2018) An efficient theta-join query processing in distributed environment. J Parallel Distrib Comput 121:42–52
11. Mittal M, Singh H, Paliwal K, Goyal L (2017) Efficient random data accessing in MapReduce. In: International conference on infocom technologies and unmanned systems (trends and future directions), ICTUS. IEEE, pp. 552–556
12. Patil A, Patil R, Pande A, Patil B (2016) Secured hadoop as a service based on infrastructure cloud computing environment. Int J Adv Res Comput Commun Eng 5(3):2319–5940
13. Tang Q, Bressoud T (2015) Analysis and modeling of mapreduce's performance on hadoop YARN. In: Proceedings of the midstates conference for undergraduate research in computer science and mathematics, MCURCSM

Factors Influencing the Selection of a Blockchain Platform for Incorporating Data Provenance into Smart Contracts

O. L. Mokalusi, R. B. Kuriakose, and H. J. Vermaak

Abstract A prominent application of blockchain smart contracts is in the modern manufacturing supply chain. This is mostly done to ensure data provenance. This research paper analyse the factors influencing the selection of a blockchain platform for incorporating data provenance into smart contracts. This research paper is written based on the literature study of various blockchain platforms. The current challenge that necessitated this research stems from the low performance, auction-based consensus mechanism, and high transaction fees of the blockchain platforms. Such contributing factors have negative financial implications associated with the deployment or interaction of smart contracts incorporating data provenance in the modern manufacturing supply chain. This research paper first gives details of some of the factors influencing the selection of smart contracts, then goes to compare how some of the latest blockchain technologies perform against these factors. Finally, analysis of the best possible blockchain platform for incorporating smart contracts is proposed.

Keywords Data provenance · Blockchain · Smart contracts · Supply chain · Smart manufacturing

1 Introduction

Blockchain is a write-only distributed database on which operates on the principle that transactions are trackable, irreversible [1] and can only be added and not edited or removed [2]. Blockchain technology enables businesses to build smart contracts, which facilitate collaborations and complex inter-organizational business processes. Smart contracts are modular, repeatable, autonomous scripts, which

O. L. Mokalusi (✉) · R. B. Kuriakose · H. J. Vermaak
Central University of Technology, Bloemfontein, Free State, South Africa
e-mail: lmokalusi@cut.ac.za

H. J. Vermaak
e-mail: hvermaak@cut.ac.za

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_47

organisations implement to automate complex transactional rules that champion blockchain principles.

A prominent application of blockchain smart contracts is in the modern manufacturing supply chain [3]. This is mostly done to ensure data provenance. Data provenance [4, 5] is defined as the place of origin or the earliest known history of a product which is tamperproof, verifiable, and ensuring integrity across the entire supply chain. Current technologies such as barcode [6] and radio frequency identification (RFID) [7] tracking have been found wanting in ensuring data provenance.

Blockchain technology with its obvious advantages is perceived as a possible solution to ensuring data provenance in the supply chain. However, there are certain factors which need to be considered while selecting the appropriate blockchain platform for incorporating data provenance. These include [8] performance, maturity, centralised or decentralised operation, and cost, to name a few.

This research paper aims to compare the factors influencing the selection of a blockchain smart contract against some of the common blockchain platforms available for incorporating data provenance in the is modern manufacturing scene. This research paper is structured that firstly an explanation of the contributing factors is done. Secondly, a comparison of the available blockchain platforms is done with respect to the contributing factors. This research paper is rounded off with an analysis and a conclusion of the best technique available currently for incorporating data provenance using blockchain smart contracts.

2 Factors Influencing the Selection of Blockchain Platform

Blockchain technology enables businesses to build smart contracts, which facilitate collaborations and complex inter-organizational business processes. Smart contracts are modular, repeatable, autonomous scripts, which organisations implement to automate complex transactional rules that champion blockchain principles such as transparency, traceability, and security which allow the proliferation of data provenance in the modern manufacturing supply chain.

However, there are contributing factors influencing the selection of a blockchain platform. This section firstly focuses on discussing some current contributing factors. These includes performance, maturity, consensus mechanism, centralised or decentralised operation, and cost, to name a few. Traditional supply chains have always been looking at ways in which operational processes can be optimised to yield greater profits in minimal time [1].

However, consensus mechanism properties [9] such as security, energy consumption, finality, throughput, and scalability are some of the contributing factors influencing the selection of a blockchain platform for incorporating data provenance into smart contracts. Consensus mechanism is an algorithm utilised to process, verify, and approve blocks with transactions by computationally resourceful miners [10] in a blockchain platform. Blockchain miners have to find the nonce [11] which is a

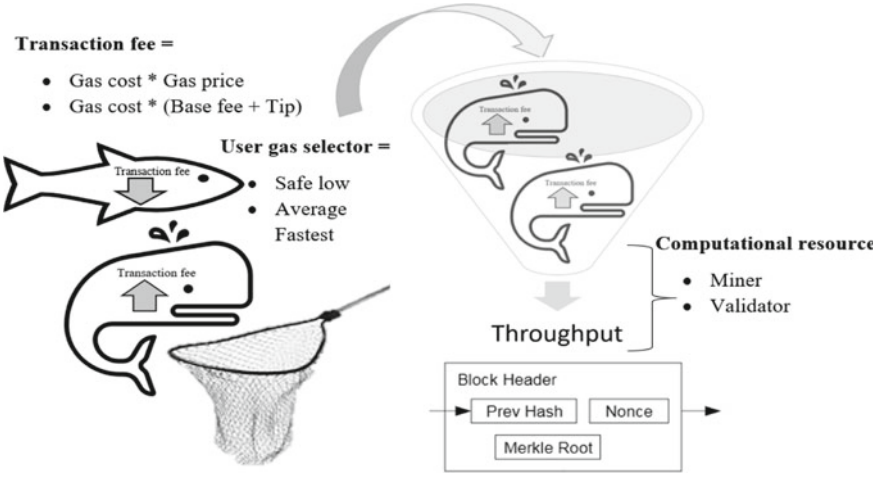


Fig. 1 Auction mechanism algorithm as a first-price auction [13–15]

number that blockchain miners have to solve for, to link the block onto the blockchain network for a transaction fee reward.

There are several existing consensus mechanisms which including and not limited to Proof of Work (PoW) and Proof of Stake (PoS) mining algorithms. Mining algorithm PoW and its fee model [12] are based on an auction mechanism as a first-price auction [13, 14].

Auction mechanism (see Fig. 1) is when miners are profiting from pickup transactions by the highest transaction fee. Users can choose between slow, average, and fast confirmation time or specifying a transaction fee manually, having an influence on performance which includes throughput and transactions latency.

Escalation of high transaction fees of the network resources has an influence in the selection of a blockchain platform for incorporating data provenance into smart contracts.

Taxonomy of blockchain-enabled smart contract-based studies [16–18] seen in Fig. 2 is mainly for improvement proposals (IPs) covering technical specification for resource-driven and optimization-driven smart contract. Other studies have been proposed for smart contract optimisation by transaction fee [18].

The development and implementation of consensus mechanism algorithms, which limit the number of miners on a network, were paramount to automating the fee bidding through a few validators with the use of PoS. The fewer the number of miners on a network means lesser the consumption of electricity [19, 20]. This would make cryptocurrency mining sustainable [21], less profitable for the miners and reduce transaction fee estimations for incorporating data provenance into smart contracts.

Consequently, such factors need to be considered while selecting the appropriate blockchain platform for incorporating data provenance into smart contracts. These

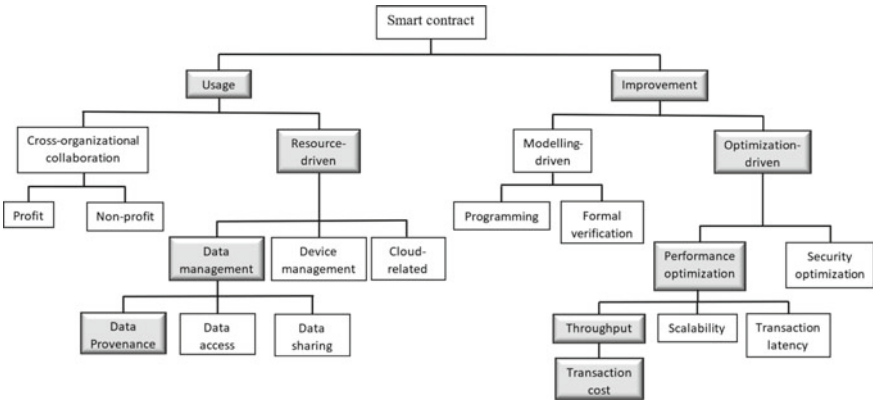


Fig. 2 Taxonomy of blockchain-enabled smart contract-based studies [16–18]

include performance, maturity, centralised or decentralised operation, and cost, to name a few.

2.1 Performance

Blockchain platforms have factors influencing their performance which includes scalability, throughput, and transactions latency [18]. Throughput is the quantity of data being sent and received within a unit of time such as Transactions Per Second (TPS). Users can choose between slow, average, and fast confirmation time or specifying a transaction fee manually having an influence on the throughput.

2.2 Maturity

Mature blockchain platforms typically have capabilities of interoperability vendor support and a stronger development community experience. Interoperability is on layer 2 [18] which offers scalability resulting in high throughput in terms of TPS. Blockchain platforms like Ethereum (legacy) have stronger development community experience which offers several tools for implementing, testing, and deploying of decentralised applications (dApps).

This would allow a smart contract incorporating data provenance to reduce transaction fees utilising transaction-inefficient code, making the process trackable and irreversible.

2.3 Cost

Cost is one of the measures implemented to avoid issues of network abuse. Computation resource executions in Ethereum are subject to a unit of gas fee [22]. Computation resources within the network are rewarded with a transaction fee which is gas price measured in "gwei" [23], for example, Ethereum gas price called Ether equivalent to 1e18 wei.

3 Blockchain Platforms

As explained in the introduction, blockchain technologies pave way for immutable transactions while ensuring transparency, traceability, and security. Some of the common available blockchain platforms include Ethereum (legacy), Ethereum 2.0, Polygon, and Tron. This section explains the salient features of each of these blockchain platforms.

3.1 Ethereum (Legacy)

Ethereum (legacy) is a smart contract blockchain platform of choice which forms layer 1 in terms of the blockchain architecture layers [24] based on its security, interoperability, and developer experience. Ethereum (legacy) currently adopts PoW as the consensus algorithm which offers low throughput, high-transaction fees, and no sovereignty [24]. Polygon is layer 2 (see Fig. 3) solution for Ethereum (legacy).

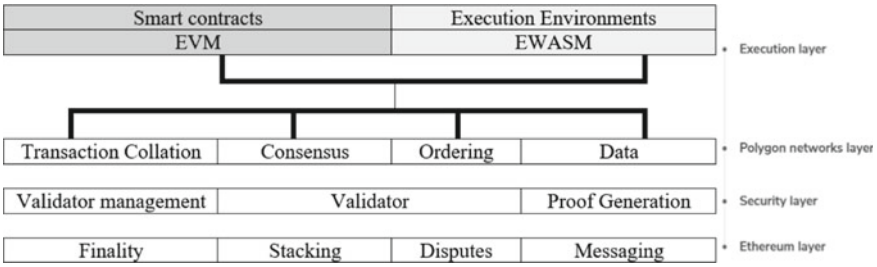


Fig. 3 Polygon architecture layers [24]

3.2 Polygon

Polygon blockchain platform enables interoperability which combines the best features of a stand-alone blockchains with sovereignty, scalability, and flexibility. Polygon currently adopts PoS as the consensus algorithm which offers high throughput, low-transaction fees, and sovereignty [24].

3.3 Ethereum 2.0

Ethereum (legacy) blockchain platform has matured to Ethereum 2.0 with the implementation of improvement proposal EIP-1559 [25] which gets rid of the first-price auction and replaces it with a fixed-price sale. Ethereum 2.0 (EIP-1559) was introduced to improve the prediction of transaction fees, to reduce throughput bottlenecks.

The recommendation for EIP-1559 [26] is divided into Phase 0, Phase 1, Phase 1.5, and Phase 2 for upgrading Ethereum (legacy) to Ethereum 2.0. Phase 0 launched the Beacon chain (test network) in December 2020 to implement PoS. PoS requires validators to stake 32 Ether. Validator stacking cannot be unlocked until Phase 1 is rolled out. Phase 1 is estimated to last for 18 months after genesis. Phase 1.5 will merge Ethereum (legacy) PoW with Ethereum 2.0 which utilises PoS. Phase 2 will add improvements to smart contract executions [27].

3.4 Tron

There are blockchain platforms that offer improved robust smart contract implementation like Tron [28]. Tron currently adopts Delegated Proof of Stake (DPoS) as the consensus algorithm, which enables a maximum of 2000 TPS [28]. DPoS blockchain platforms are generally faster with higher per second transaction rates than blockchain platforms with consensus mechanism like PoW and PoS.

However, DPoS is in its infancy, and it is not generally considered secure enough to be the utilised in money-transacting blockchain platforms.

4 Factors Influencing the Selection of a Blockchain Platform

A comparison of blockchain platforms such as Ethereum (legacy), Ethereum 2, Polygon, Tron (see Table 1) is conducted with respect to factors influencing the

Table 1 Comparison of blockchain platforms with respect to factors influencing the selection of a blockchain platform [11, 13, 23, 24, 26, 28–31]

	Ethereum (legacy)	Ethereum 2.0	Polygon	Tron
Currency and symbol	ETH	ETH	MATIC	TRON
Cost	High [31]	Low [23]	Low [30]	Low [28]
Performance	<5 min [31]	High [26]	30 s [30]	3 s [28]
TPS	Low [31]	High [13]	High [24]	High [28]
Throughput	Low [26]	High [29]	High [24]	High [30]
Consensus algorithm	PoW [11]	PoS [29]	PoS [24]	DPoS [28]
Cheapest transfer fee	$3.2606 [31]	$0 [26]	$0 [30]	$0 [28]

selection of a blockchain platform for incorporating data provenance into smart contracts.

5 Analysis

With the infinite of blockchain platforms, this analysis explores factors influencing the selection of a blockchain platform for incorporating data provenance into smart contracts. Firstly, the comparison of blockchain platforms will be limited to the following blockchain platforms Ethereum (legacy), Ethereum 2.0, Polygon, Tron (see Table 1).

With respect to cost, Tron and Polygon blockchain platforms are lower than Ethereum (legacy). Ethereum 2.0 is still in Phase 0 which is mainly a test network (testnet). Testnet transaction fees are not a considered since transaction is free and testnets do not experience network congestion like the public main network (mainnet). However, Ethereum on full implementation will offer no first-price auction and replace it with a fixed-price sale which will be result in low transaction fees.

With respect to performance, Tron blockchain platform is higher than Ethereum (legacy). With respect to throughput, Ethereum 2.0, Polygon, and Tron blockchain platforms are higher than Ethereum (legacy). With respect to consensus algorithm, Tron blockchain platform offers the preferred DPoS which is better than Ethereum (legacy) PoW which is an inefficient auction mechanism. With respect to cheapest transaction fees, Ethereum 2.0, Polygon, and Tron blockchain platforms are low.

Table 2 Experiments proposed to be conducted on the mainnet

Blockchain platform	Small contract	Large contract
Ethereum (legacy)	Mainnet	Mainnet
Ethereum 2.0	Testnet	Testnet
Polygon	Mainnet	Mainnet
Tron	Mainnet	Mainnet

However, Tron remains outstanding as it offers higher performance followed by Polygon.

6 Future Work and Conclusion

Consensus mechanism such as PoW mining algorithm has a major contributing factor on security and selection of a blockchain platform. Organisations switch from the PoW algorithm to new consensus mechanisms that promise lower transaction fees as well as lower energy costs the like Tron and Polygon for incorporating data provenance into smart contracts. Traditional supply chains have always been looking at ways in which operational processes can be optimised to yield greater profits in minimal time.

Based on the analysis in section 5, the results point to Tron and Polygon as the blockchain platforms of choice with factors such as low transaction fees, high performance, high throughput, and the utilisation of efficient PoS consensus algorithm for incorporating data provenance into smart contracts.

However, future experiments are proposed and will be conducted on the mainnet and testnet as outlined in Table 2. Factors influencing the selection of a blockchain platform such as transaction fees, performance, and throughput will be tested to confirmed and validate the analysis made in this paper for incorporating data provenance into smart contracts. A blockchain platform which is most cost effective and offers higher performance will be confirmed.

References

1. Prause G (2019) Smart contracts for smart supply chains. IFAC-PapersOnLine 52:2501–2506. https://doi.org/10.1016/j.ifacol.2019.11.582
2. Cui P, Dixon J, Guin U, Dimase D (2019) A blockchain-based framework for supply chain provenance. IEEE Access 7:157113–157125. https://doi.org/10.1109/ACCESS.2019.2949951
3. Mokalusi OL, Kuriakose RB, Vermaak HJ (2022) Exploring the means and benefits of including blockchain smart contracts to a smart manufacturing environment: water bottling plant case study. pp 309–319. https://doi.org/10.1007/978-981-16-6369-7_27
4. Jaquet-Chiffelle DO, Casey E, Bourquenoud J (2020) Tamperproof timestamped provenance ledger using blockchain technology. Forensic Sci Int: Digital Invest 33:300977. https://doi.org/10.1016/j.fsidi.2020.300977
5. Demestichas K, Peppes N, Alexakis T, Adamopoulou E (2020) Blockchain in agriculture traceability systems: a review. Appl Sci (Switzerland) 10:1–22. https://doi.org/10.3390/APP10124113
6. Jardine N, Gericke GA, Kuriakose RR, Vermaak HJ (2019) Wireless SMART product tracking using radio frequency identification. In: 019 IEEE 2nd wireless Africa conference, WAC 2019—Proceedings. https://doi.org/10.1109/AFRICA.2019.8843418
7. Duroc Y, Tedjini S (2018) RFID: a key technology for Humanity. C R Phys 19:64–71. https://doi.org/10.1016/j.crhy.2018.01.003

8. Top 9 blockchain platforms to consider in 2021. https://searchcio.techtarget.com/feature/Top-9-blockchain-platforms-to-consider
9. Oyinloye DP, Teh JS, Jamil N, Alawida M (2021) Blockchain consensus: an overview of alternative protocols
10. Ren W, Hu J, Zhu T, Ren Y, Choo KKR (2020) A flexible method to defend against computationally resourceful miners in blockchain proof of work. Inf Sci 507:161–171. https://doi.org/10.1016/J.INS.2019.08.031
11. Wood G (2016) Ethereum: a secure decentralised transaction ledger. EIP-150 revision
12. Gas and fees. https://ethereum.org/en/developers/docs/gas/
13. EIP 1559 and Fee Structure
14. Leonardos S, Monnot B, Reijsbergen D, Skoulakis S, Piliouras G (2021) Dynamical analysis of the EIP-1559 ethereum fee market
15. Pierro GA, Rocha H (2019) The influence factors on ethereum transaction fees. In: 2019 IEEE/ACM 2nd international workshop on emerging trends in software engineering for blockchain (WETSEB). pp 24–31
16. Tasca P, Tessone CJ (2019) A taxonomy of blockchain technologies: principles of identification and classification. Ledger 4. https://doi.org/10.5195/ledger.2019.140
17. Labazova O, Dehling T, Sunyaev A (2019) From hype to reality: a taxonomy of blockchain applications. In: Proceedings of the annual Hawaii international conference on system sciences. IEEE Computer Society, pp 4555–4564
18. Khan SN, Loukil F, Ghedira-Guegan C, Benkhelifa E, Bani-Hani A (2021) Blockchain smart contracts: applications, challenges, and future trends. Peer-to-Peer Netw Appl 14:2901–2925. https://doi.org/10.1007/s12083-021-01127-0
19. Cambridge Bitcoin Electricity Consumption Index (CBECI). https://cbeci.org/index
20. Bitcoin consumes "more electricity than Argentina"—BBC News. https://www.bbc.com/news/technology-56012952
21. Náñez Alonso SL, Jorge-vázquez J, Echarte Fernández MÁ, Reier Forradellas RF (2021) Cryptocurrency mining from an economic and environmental perspective. analysis of the most and least sustainable countries. Energies 14. https://doi.org/10.3390/en14144254
22. What is Gwei?—ETH Gas Station. https://legacy.ethgasstation.info/blog/gwei/
23. What Is Gas (Ethereum)? https://www.thebalance.com/gas-ethereum-5189030
24. Polygon. https://polygon.technology/
25. Watch The Burn: EIP-1559 real-time ETH burn visualization for ethereum. https://watchtheburn.com/
26. The Unveiling of Ethereum 2.0—ETH Gas Station. https://legacy.ethgasstation.info/blog/the-unveiling-of-ethereum-2-0/
27. The Eth2 upgrades. https://ethereum.org/en/eth2/
28. TRON Ecosystem Introduction. https://developers.tron.network/docs/getting-started
29. Leonardos S, Monnot B, Reijsbergen D, Skoulakis S, Piliouras G Dynamical analysis of the EIP-1559 ethereum fee market
30. Polygon Gas Tracker. https://polygonscan.com/gastracker/
31. ETH Gas Station. https://legacy.ethgasstation.info/

Improving Model Accuracy by Means of Explanations

Daiki Yamaguchi, Israel Mendonça, and Masayoshi Aritsugi

Abstract Recently, artificial intelligence (AI) methods are becoming popular due to their efficacy when applied to different fields; hence, there has been an increasing effort to provide meaningful explanations of the decision-making process of such methods. For computer vision methods, one of the most utilized is called layer-wise relevance propagation (LRP), which computes the relevance of each layer in the process of classification and generates a comprehensible relevance map image representing the importance of each part of the image during the classification process. In this work, we take one step toward the utilization of explainable artificial intelligence (XAI) methods, such as LRP, to improve the accuracy of well-established AI models. We propose a novel methodology that creates a meta-model that is able to learn the classification process of a model through observation of LRP relevance map images. This meta-model is able to improve the accuracy of the baseline model in up to 30% in the best case and 1% on average.

Keywords Machine Learning · Explainable AI · Deep learning · Meta-models · Image classification

1 Introduction

Deep learning (DL) techniques have been successfully employed in computer vision problems. It is expected that it will play an active role in various fields becoming an indispensable tool for achieving success. However, DL has an issue that still needs to be addressed, namely explainability [1]. DL algorithms have a black-box nature, where the basis for decisions is unknown due to the complexity of the calculations

D. Yamaguchi · I. Mendonça (✉) · M. Aritsugi
Kumamoto University, Kumamoto, Japan
e-mail: israel@cs.kumamoto-u.ac.jp

D. Yamaguchi
e-mail: d-yamaguchi@st.cs.kumamoto-u.ac.jp

M. Aritsugi
e-mail: aritsugi@cs.kumamoto-u.ac.jp

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_48

involved; hence, it is hard to reason about the decision process that led a DL to a determined output.

The field of explainable artificial intelligence (XAI) tries to mitigate this problem by providing techniques that gives simple human-comprehensible explanations. One of the main XAI techniques is called layer-wise relevance propagation (LRP) [2]. It calculates the relevance of the input to the output and produces a relevance map for the input. This relevance map resembles the original image; however, it highlights the areas of the image in which the model used for classification.

The main purpose of this relevance map is for explaining the basis of the AI's decision. It is a non-invasive technique, in the sense that it does not affect the model from which it is explained or the output in any way. However, the relevance map generated by LRP from the model captures characteristics of the input very well.

Since this relevance map is able to capture the importance of the input's features well, it is natural to think that a model that is trained using such relevance maps as inputs would be able to have a higher accuracy than the base model.

In this paper, we demonstrate that a meta-model trained using relevance map obtained from a baseline deep learning network is able to outperform the baseline by up to 30%. Also, we show that a combination of the baseline and this meta-model creates an even stronger classifier that is able to achieve even higher accuracy.

In short, the contributions of this research are as follows:

- We take one step toward understanding how XAI techniques can be used for purposes other than explainability.
- Through a series of experiments, we show the importance of using the relevance maps to achieve high-performance models.
- We introduce a technique that is model agnostic, and it is empirically shown that our technique could improve the accuracy of deep learning networks.

The rest of this paper is organized as follows: Sect. 2 gives a brief introduction of the mechanics of LRP. Section 3 explains our technique; Sect. 4 describes our experimental methods and results. Section 5 concludes the discussion and explains points of improvement, and finally, Sect. 6 explains points of improvement in this work and gives some directions on what we could do to improve such points.

2 Background

Layer-wise relevance propagation (LRP) [2] is a method to obtain the relevance R of an input x to an output of a model f. When the input x is an image, a heat map can be created based on the relevance R to show where the model focuses its attention in the image. Figure 1 explains how to apply LRP to a neural network (NN).

In this NN, the output of neuron j, x_j has a nonlinear activation function of σ and a bias term of b. This can be expressed as:

$$x_j = \sigma(\Sigma_i w_{ij} x_i + b). \tag{1}$$

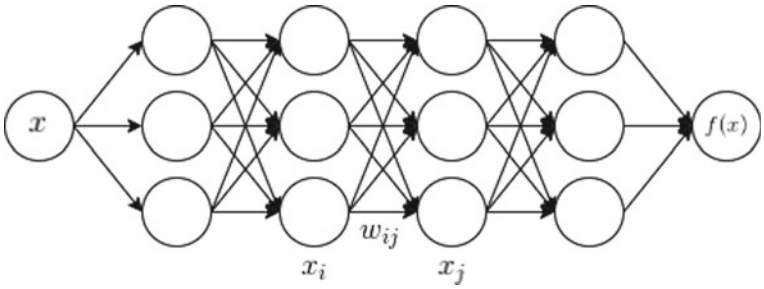


Fig. 1 Propagation of neural network: If there is a model and given a input, the output of a neuron is related to the weight of each layer multiplied by the output of the previous layer

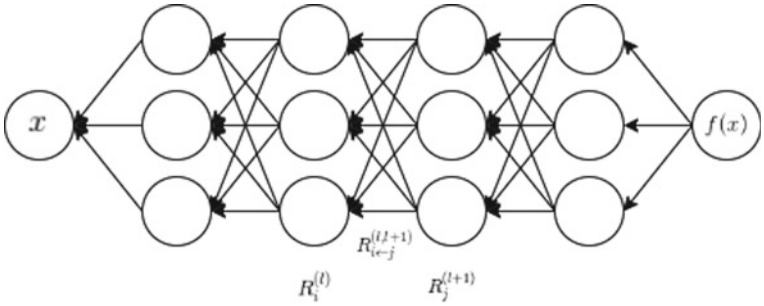


Fig. 2 Layer-wise Relevance Propagation: The relevance of a neuron is related to the relevance of each layer multiplied by the message of the previous layer. Sum of the relevances of the same layer is conserved

In LRP, when there is a combination of input x and output $f(x)$, the relevance is propagated backward from the output layer to the input layer, as shown in Fig. 2.

If the relevance $R_j^{(l+1)}$ of neuron j in the $(l+1)$ layer is transmitted to the relevance $R_i^{(l)}$ of neuron i in the l layer as $R_{i\leftarrow j}^{(l,l+1)}$, then this can be expressed as:

$$R_j^{(l+1)} = \Sigma_i R_{i\leftarrow j}^{(l,l+1)}. \tag{2}$$

In [2], the authors proposed a formula to calculate the relevance $R_{i\leftarrow j}^{(l,l+1)}$ using the ϵ-rule, represented as:

$$R_{i\leftarrow j}^{(l,l+1)} = \frac{z_{ij}}{z_j + \epsilon \cdot sign(z_j)} \cdot R_j^{(l+1)}. \tag{3}$$

Here, $z_{ij} = x_i w_{ij}$ and $z_j = \Sigma_i z_{ij}$.

The propagation rules of the transmission $R_{i\leftarrow j}^{(l,l+1)}$, including those not shown here, are the sum of the relevance of each layer is conserved. In other words, since $\Sigma_i R_i^{(0)} = \cdots = \Sigma_i R_i^{(l)} = \Sigma_i R_i^{(l+1)} = \cdots = f(x)$, the relation $\Sigma_i R_i^{(0)} \approx f(x)$ holds.

3 Methodology

In this section, we explain how we created a meta-model that utilizes LRP images to increase baseline model performance. More specifically, Sect. 3.1 explains the meta-model creation process in details, while Sect. 3.2 explains the process of the creation of the data for our meta-model. Lastly, Sect. 3.3 explains how we utilize this meta-model for prediction.

3.1 Meta-Model

Our principal aim is to improve the accuracy of a baseline model without needing to modify any of its architecture nor weights. To that end, we considered a machine learning technique that is able to learn the classification pattern of a baseline model and then use this pattern to generate better predictions. We expect that our meta-model can improve the accuracy of the baseline model via effective classification pattern recognition. Unlike traditional machine learning models, our meta-model is optimized to fit the observed base model relevance pattern for each class, rather than the shape of the original image itself. Empirically, this can significantly improve accuracy, specially over low-accuracy datasets.

Our methodology works as follows. Consider a base model m_{base} trained for a dataset D of images. For each single image d_i in D, we apply the LRP algorithm to m_{base}. The output of this operation is an image d_{iLRP}, which is stored in a secondary dataset D_{LRP} that contains the LRP version of each image in D. Our meta-model, which is of the same architecture of the base model, is trained by using the set D_{LRP} as training data instead of the original dataset D. This procedure is illustrated in Fig. 3.

By learning from the base model relevance recognition pattern instead of the original shape, our meta-model is able to learn samples in which the base model miss-classify the input image. Since the shape or the original image is also contained in the LRP images, we can consider that the D_{LRP} dataset is an augmented set of D, in the sense that it contains not only the original image shape, but also some meta-information on how the baseline model sees it.

3.2 Meta-Model Dataset Generation

As explained in Sect. 3.1, our meta-model is trained using LRP's relevance map images as input. However, for each image d_i in the dataset D, LRP generates $k \in K$ different images, in which K is the number of classes in the dataset. This leads to the discussion on how to use the images in the training process.

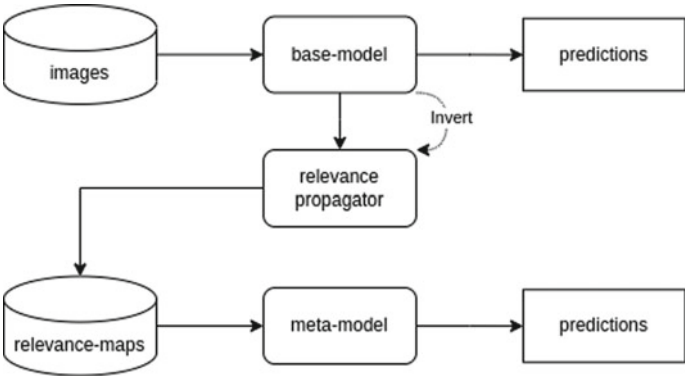


Fig. 3 Relationship between the base model and the meta-model. The base model learns to classify the images and makes predictions. The relevance propagator is created by the inverted image model and generates relevance maps. The meta-model is trained in these maps and makes predictions

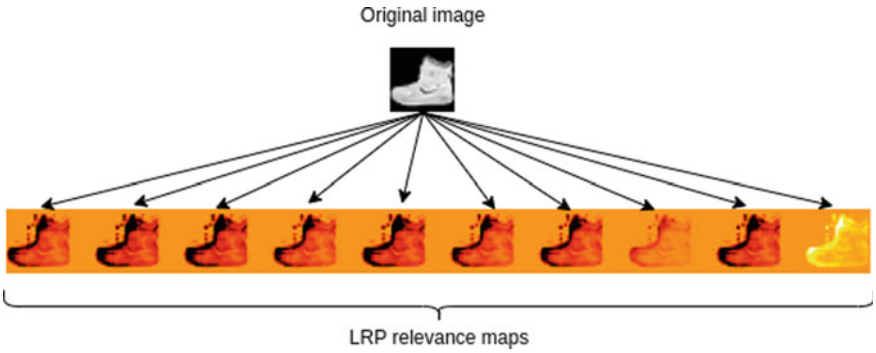


Fig. 4 Example of one LRP image being created from the original dataset. The Fashion-MNIST is a dataset containing 10 classes, and for each class, a relevance image is created and aligned. In this example, the correct label for the figure is 9 (Ankle boot)

The most straightforward way to use the images is to align them as a $k - dimensional$ image. However, this approach would require modifications in the model head since the input would increase from a $3 - dimensional$ image (RGB) to a $3 \times k - dimensional$ image, and as result, it would become unpractical. Another approach is to combine the $k - dimensional$ images into a single image. This approach is easy to implement since it only requires a simple image combination procedure and does not require any modifications in the model heads.

In our work, we align each LRP image side-by-side generating, ordered by the class number from left to right. Figure 4 illustrates the creation process of our LRP image using a single image of the Fashion-MNIST [3] dataset.

3.3 Training and Prediction

Using information provided by Sects. 3.1 and 3.2, we can divide our training procedure in three steps.

The first step starts by training the baseline model normally. After the baseline model is trained, Step 2 starts by generating the LRP images. The process consists of passing each image through the network and storing the generated relevance image on disk. On Step 3, we use the images generated on Step 2 to train our meta-model. It is important to say that the original train images utilized on the training on the baseline model are not utilized on the training of the meta-model but only the LRP relevance map images. After all the three steps are completed, we have two models: the original baseline model and the meta-model.

For prediction, we tried two approaches: The first one, as illustrated in Fig. 5a, is a straightforward classification procedure that an LRP relevance map image is generated by passing an input image through the base model. Then, this relevance map image is used as input in our meta-model. The meta-model outputs a class probability vector, in which each entry k is the probability of the image belonging to the class k. This procedure is similar to the classification using the base model, but instead of the original image being used as input, we use an LRP relevance map image instead.

The second approach for prediction is an ensemble. In this approach, we use not only the LRP relevance map generated by the base model but its predictions as well. As we can see in Fig. 5b, similarly to the first prediction approach, we feed the LRP relevance maps to the meta-model; however, when it comes to calculate the predictions, we do a pondered average of the predictions of the base model and the meta-model. The weight of each model can range from 0 to 1, as long as the sum of the weights is 1. The higher the weight of a model, the higher the emphasis on the predictions of the model.

4 Experiments

In this section, we quantitatively verify the effect of utilizing our meta-model method alongside the baseline model. We compare the performance gains in terms of accuracy.

We used four different common computer vision datasets in which the details can be confirmed in Table 1. For our baseline model and the meta-model, we used ResNet-18 with pretrained weights. For the baseline model, we used “moderate data augmentation” (flip/translation), and training is done using stochastic gradient descent for 50 epochs, and patience of 10. All the models were trained using the PyTorch [4] library. For LRP, we created a custom implementation based on the original paper. The performance metric for all models is accuracy, in which we calculate how many times a model classifies correctly the class of a given image.

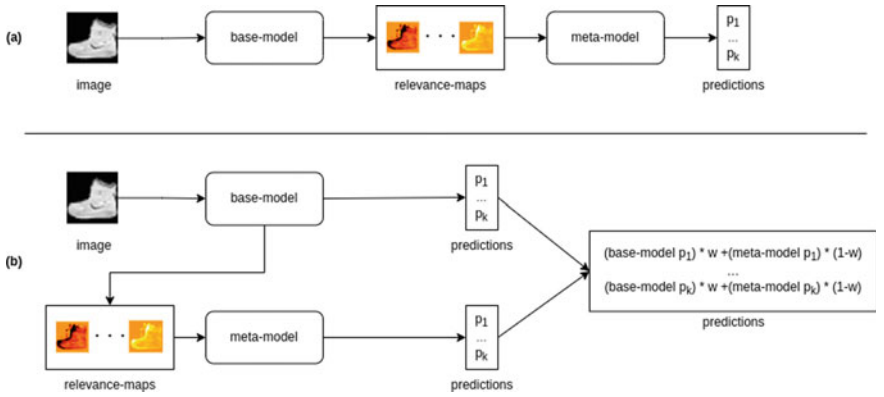


Fig. 5 Prediction approaches: **a** We use the predictions generated solely from the meta-model; **b** We use the predictions from the original baseline model together with the meta-model

Table 1 Datasets details

Dataset name	Train size	Test size	#classes
Fashion-MNIST [3]	60,000	10,000	10
MNIST-BG [5]	60,000	12,000	10
MNIST-BG-Rotated [6]	60,000	12,000	10
CIFAR10 [7]	50,000	10,000	10

The above mentioned datasets were chosen because we wanted to test our methodology in different scenarios. Fashion-MNIST can be considered an easy dataset, in the sense that accuracy for this dataset is usually high. We tested if our meta-model was able to increase the accuracy of an already high dataset. MNIST background (mnist-bg) and MNIST background rotated (mnist-bg-rotated) are harder datasets (when compared to Fashion-MNIST) due to the amount of noise and rotation applied to the images. On those datasets, we wanted to verify if our meta-model was able to capture a pattern of mistakes from the baseline model. Lastly, since the previous introduced datasets were on grayscale, we wanted to verify if our meta-model was able to improve on RGB images as well.

4.1 Results

4.1.1 General Comparison

We start our experiments by evaluating the performance of the base model against our meta-model, and the ensemble of base model and meta-model. These are the equivalent of prediction style 1 (P1) and prediction style 2 (P2) as explained in

Table 2 Accuracy percentage of each model per dataset

Dataset	Baseline-model	Meta-model (P1)	Ensemble (P2)
Fashion-MNIST	0.9241	0.9273	**0.9275**
MNIST-BG	0.6136	**0.8126**	0.8122
MNIST-BG-Rotated	0.5356	0.5585	**0.5705**
CIFAR10	0.8185	0.8156	**0.8187**

Sect. 3.3. The results can be verified at Table 2. For the ensemble in P2, we are using $weight = 0.5$, i.e., we average the probabilities of the base model and the meta-model.

By using only our meta-model, we are able to increase the accuracy for 3 out of 4 datasets. It efficiently uses the relevance maps generated by the base model and improves the accuracy based on the information contained on it. Among the improved datasets, the one with the higher improvement in accuracy was MNIST-BG. We can say that the meta-model was able to identify when the base model was seeing noise and improved the accuracy for around 30% of the base model. The second big improvement was on MNIST-BG-Rotated, which was the hardest dataset. Still, our meta-model increased the baseline accuracy in around 4%. For MNIST, the improvement was smaller around 0.03%. The only dataset in which our meta-model fails to improve the accuracy CIFAR10.

One of the main differences of using our methodology on grayscale images and colored ones is that the LRP images do not carry any information on different colors. While grayscale image classification relies on shape, RGB images use colors to get information. When we use LRP, the color information is lost although the relevance focus point is still kept in relevance map. However, there is a difference between focus point highlight and actual color difference.

As we can see in Fig. 6, two points of interest that are separated by color can become merged into a single relevance point when we mapped from the original image to the LRP relevance map. We believe this is the main reason that performance is hurt when using RGB images instead of gray scale.

For the experiments using ensemble, it can be verified that for all cases its results are better than the baseline. Ensembling models is usually a good approach to improve a model performance because it adds diversity to the model. Apart from MNIST-BG, the ensemble of our meta-model together with the baseline model was able to have better results than each of them separated.

4.1.2 Ensemble Weight Distribution

In this part of our experiments, we verify the impact of using different weights for the ensemble of the base model and meta-model. As explained before, the ensemble uses a pondered average on the predictions of both models to generate a final prediction.

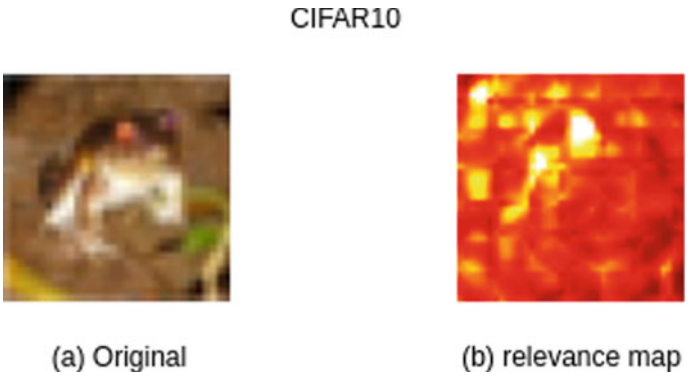


Fig. 6 RGB image difference: **a** Original image in which the boundary of each part of the image is defined by color; **b** Color information is lost during the conversion part as only the relevance is colored

Table 3 Model results per dataset

Dataset	Baseline	MM 0.2	MM 0.4	MM 0.5	MM 0.6	MM 0.8
Fashion-MNIST	0.9241	**0.9277**	0.9276	0.9275	0.9273	**0.9277**
MNIST-BG	0.6136	0.7841	0.8092	0.8122	0.8136	**0.8139**
MNIST-BG-Rotated	0.5356	**0.5749**	0.5736	0.5705	0.5679	0.5624
CIFAR10	0.8185	0.8188	**0.8193**	0.8187	0.8185	0.8183

We vary the meta-model importance by setting the weight as 0.2, 0.4, 0.5, 0.6, and 0.8 and compare the impact it has in accuracy. Notice that a weight of 0.2 means that the base model prediction will have 80% importance, while the meta-model will have only 20% importance. The impact of these differences can be verified at Table 3

The results show that any combination of the base model with the meta-model outperforms all the classification performances of the base model alone. This supports the idea that using the meta-model as support for the base model can bring benefits to accuracy. However, despite both models benefiting from being combined with each other, there are differences as big as 3% in the accuracy when using different weights.

4.2 Discussion

From the results of our experiments, we found that the combination of the base model and the meta-model surpassed the results of the base model. This suggests that the meta-model may be able to capture features that are not captured by the base model.

In addition, Table 3 shows that the combination of the results of the base model with the results of the meta-model surpasses the results of the base model. However, since the weights of the best combination differ depending on the dataset, it is necessary to calculate the weights after training the model.

5 Conclusion

In this paper, we introduced a method to improve the classification performance of a base model by using a meta-model that uses LRP relevance image maps as training data. We outperformed the base model in both test cases: using our meta-model only and using as ensemble with the base model. We also performed a study on the importance of weight distribution on the ensemble and how the meta-model helps improving accuracy. We also studied the case in which our methodology may not have much impact.

6 Future Work

Despite yielding good results, there is still room for improvement in the way explanations are used. In this paper, we used two different models to achieve the results; however, a novel architecture in which these models are connected, and we can do a one-shot prediction is preferable. Also, our model learns from an image that is a single concatenation of the relevance map of each class. For a small number of classes as the ones tested in this paper, our approach is feasible; however, as the number of classes increases, the size of the training image can become unpractical for training and predicting. In that case, a technique to select only a portion of the class images may be preferable in order to reduce the input image size. This selection method would have to be studied.

References

1. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A,Tabik S, Barbado A, García S, Gil-López S, Molina D, Benjamins R et al (2020) Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. In: Information fusion, vol 58, pp 82–115
2. Bach S, Binder A, Montavon G, Klauschen F, Müller K-R, Samek W (2015) On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PloS one 10(7):0130140
3. Xiao H, Rasul K, Vollgraf R (2017) Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms. arXiv preprint arXiv:1708.07747
4. Pytorch. https://pytorch.org

5. MNIST + background images. http://www.iro.umontreal.ca/~lisa/icml2007data/mnist_background_images.zip
6. Rotated MNIST digits + background images. http://www.iro.umontreal.ca/lisa/icml2007data/mnist_rotation_back_image_new.zip
7. Krizhevsky A (2009) Learning multiple layers of features from tiny images. Maters's thesis, Toronto University

Analysis of an Independent Double Boost Interleaved Converter in a Renewable Energy Application

Vasile Mihai Suciu, Lucian Nicolae Pintilie, Sorin Ionuț Salcu, Petre Dorel Teodosescu, Teodor Pana, and Zsolt Mathe

Abstract In this paper a brief analysis has been carried out targeting a DC-DC step-up converter used within renewable energy applications. The converter topology involved is called as an "independent double boost interleaved converter" (IDBIC) with three-level output. The operation mode of the proposed power electronics conversion stage has been analyzed in the context of a photovoltaic application. The converter is used (in the present application) as an interfacing stage between three PV modules connected in series and the deserved load. An MPPT algorithm was also implemented as the control law that drives the application, in order to extract the maximum power from the solar panels array. A series of tests were performed in two study cases: In first case an off-line co-simulation using Plexim—PLECS and MATLAB—Simulink has been done. In second case an on-line Real-Time, Rapid Control Prototyping (RCP) simulation using C2000 F28069M Texas Instruments DSP with HOTLINK jTAG connection to the host computer that runs Altair/Solid Thinking Embed (formerly known as VisSim) as RCP development environment has been made. The resulting current and voltage waveforms for different cases of solar irradiation values were represented and analyzed. Also, the efficiency of the power stage was determined at different levels of solar irradiation.

V. M. Suciu (✉) · L. N. Pintilie · S. I. Salcu · P. D. Teodosescu · T. Pana
Technical University of Cluj-Napoca, 400489 Cluj-Napoca, Romania
e-mail: mihai.suciu@emd.utcluj.ro

L. N. Pintilie
e-mail: lucian.pintilie@emd.utcluj.ro

S. I. Salcu
e-mail: sorin.salcu@emd.utcluj.ro

P. D. Teodosescu
e-mail: petre.teodosescu@emd.utcluj.ro

T. Pana
e-mail: teodor.pana@emd.utcluj.ro

Z. Mathe
SC Tehnologistic SRL, Cluj-Napoca, Romania
e-mail: zsolt.mathe@tehnologistic.ro

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_49

Keywords DC-DC converters · Independent double boost interleaved converter (IDBIC) · Renewable energy · Maximum power point tracking (MPPT)

1 Introduction

The use of alternative—renewable energy sources in the actual context of energy pricing crisis has been proved as a significant solution—response to the energy deficit [1]. Strategies such as implementing distributed generation by using solar, wind or other renewable sources in microgrid systems [2] are proposed in order to mitigate useless energy loss and to ensure better energy production in response to the energy balance demand. Among all other sustainable energy sources, solar photovoltaic (PV) systems, are progressively becoming the most popular and affordable alternative "green" sources for electrical power generation [3, 4], being suited for distributed power generation and microgrids [5]. To integrate the PV energy production solution, specific power converters (DC to DC or DC to AC) are required [6]. Those power conversion stages can improve the overall efficiency of the system by tracking the maximum power produced by the PV array or by a single panel and meeting the voltage demands for the energy conversion stage toward the storage elements or end-user. As an example, the implementation of a master high voltage DC (HVDC) bus is ubiquitous when dealing with a PV system. In this case a high gain DC-DC boost converter allows each PV module [7] or a combination of several series or parallel modules, to match the HVDC bus voltage requirements. This is necessary due to the variable power generated and voltage fluctuations of the PV system. The proposed converter topology that has been used in this paper as an interfacing element for a PV array, was presented in [8] with its working principle, its math analysis and key features. The high gain, low voltage stress and independent control of the integrated electronic structure of the topology, should make it an appealing solution as a PV optimizer, or a DC-DC stage for a microinverter.

In the present work a brief analysis of an interleaved converter based on a patent application [9], working in a renewable energy application has been carried on. The behavioral study of the power conversion stage has been done through the simulation of the converter and its internal control law (MPPT algorithm) within different simulation environments and automated testing means and methods. The practical implementation was based on a rapid control prototyping system composed of C2000 F28069M Texas Instruments Digital Signal Processor (DSP) with Real-Time jTAG connection to a host PC with a running host application of Altair Embed (formerly known as Solid Thinking Embed or Vis Sim) [10]. The Rapid Control Prototyping Real-Time Computing System (the DSP) is used in order to execute the control law (MPPT algorithm) of the power stage, in real-life conditions, based on real signals coming from current and voltage sensors. The main advantage of the Rapid Control Prototyping Real-Time simulation consists of in the fact that fine adjustments of

the control law parameters can be performed "on the go" during the "live" or "online" simulation. The key point of the current study is to analyze the behavior of the proposed power stage in the Maximum Power Point Tracking operation mode.

Apart from the introduction, this paper is organized in the following structure. In Sect. 2 the simulation and experimental models are presented. Section 3 presents the simulation and experimental results. The conclusions are presented in Sect. 4.

2 System Setup

2.1 Simulation Model

Figure 1 represents the block diagram related to the simulation model developed in MATLAB—Simulink in conjunction with the PLECS Blockset. By using the PLECS Blockset (toolbox) within MATLAB—Simulink, a co-simulation approach of the control law and the power stage can be done. The control law has been implemented using MATLAB—Simulink's native toolset and the power stage has been implemented using the PLECS Blockset add-on.

Three PV modules connected in series have been used within the simulation model. To extract the maximum power from the PV array, the Maximum Power Point Tracking control law consisting of the Perturb and Observe algorithm has been implemented in MATLAB—Simulink. A flowchart for the implemented MPPT algorithm is presented in Fig. 8 from [11]. By using the simulation model, the power stage behavior can be recorded and analyzed in the conditions produced by the solar PV array at a specific irradiance variation profile across the three series connected panels. In Fig. 2 the simulation model implemented using MATLAB—Simulink is shown. The model consists of the PV array block emulating the characteristics of three "Canadian Solar CS6X-300P" panels connected in series, the MPPT control law algorithm sub-system and the IDBIC power stage sub-system block implemented in PLECS Blockset.

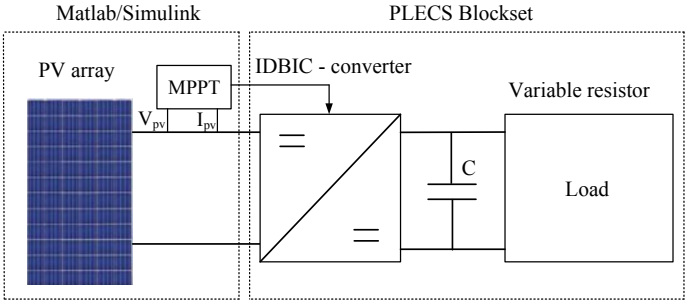


Fig. 1 Block diagram of the simulation model

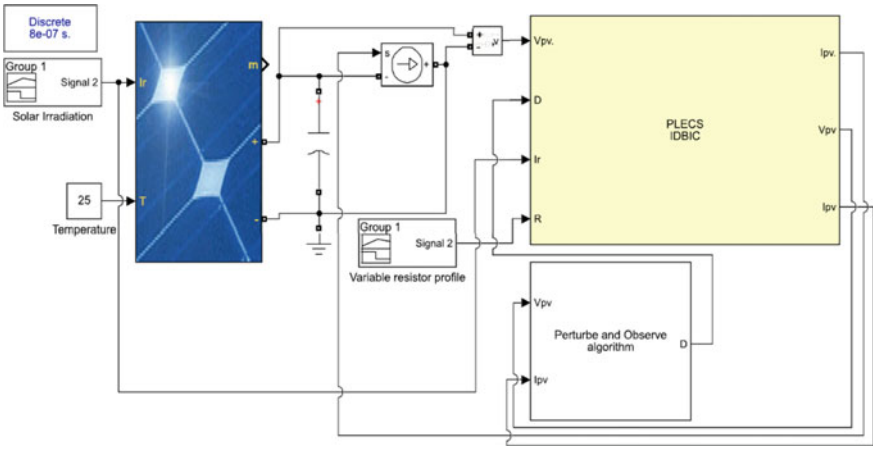


Fig. 2 Simulink block-model of the simulation model

Figures 3 and 4 are representing the IDBIC power stage modeled in PLECS Blockset and the Perturb and Observe algorithm model diagram implemented in the MATLAB—Simulink environment.

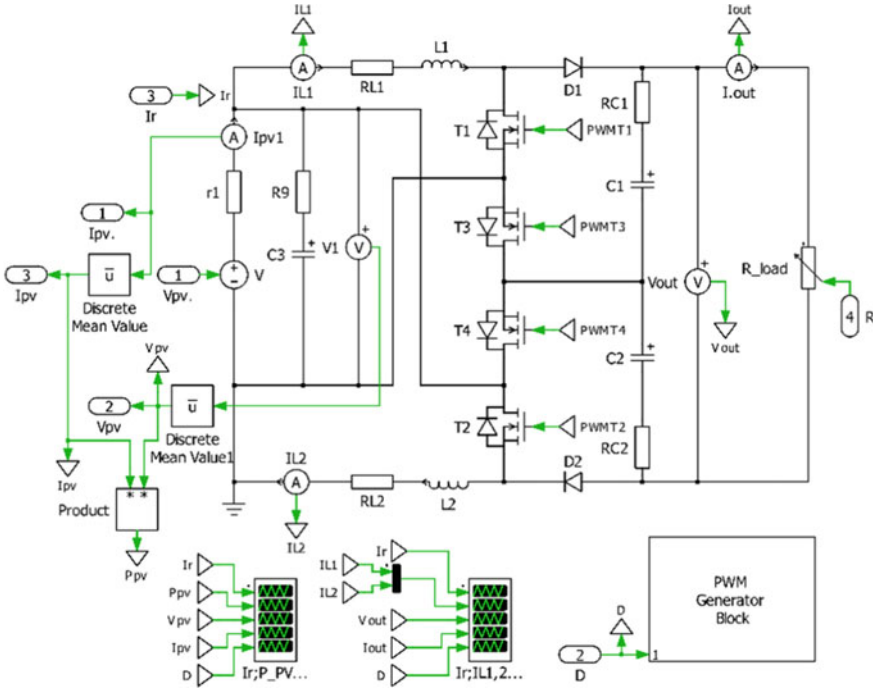


Fig. 3 IDBIC power stage modeled in PLECS Blockset

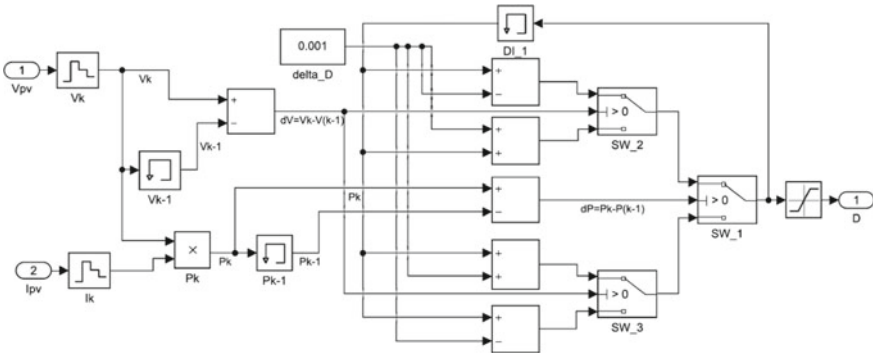


Fig. 4 Perturb and Observe control law block-model in MATLAB—Simulink

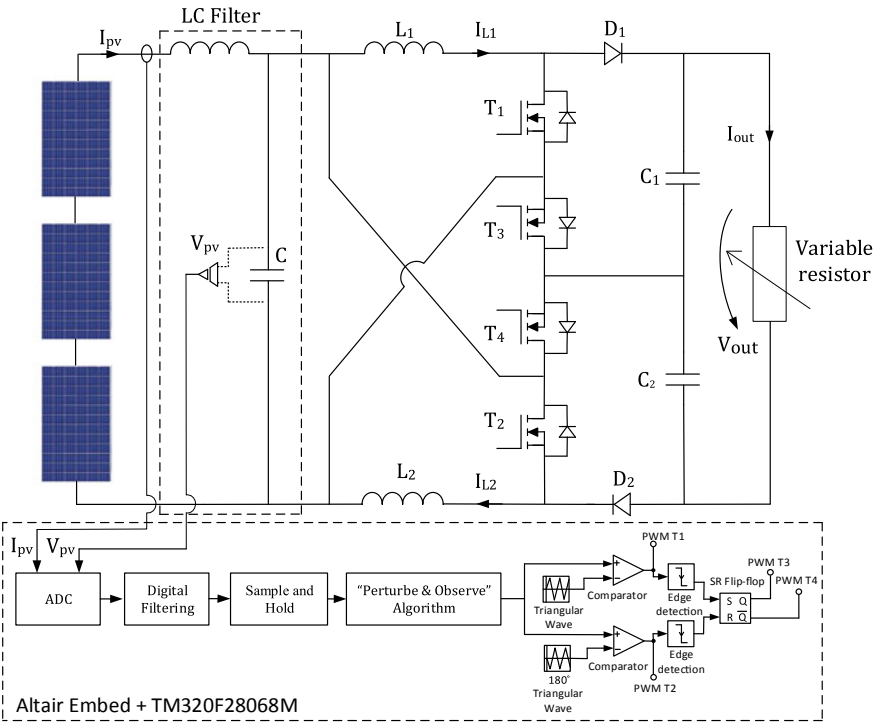


Fig. 5 Block-model of the experimental setup

2.2 Experimental Model

Figure 5 represents the electronic diagram of the experimental model and in Fig. 6 the laboratory test setup is shown. The characteristics of three "Canadian Solar CS6X-300P" solar panels connected in series were emulated by using the bi-directional power supply (with solar panel emulation capabilities) EA-PSB 9080–360 3U. The deserved load consists of a variable resistor. The equivalent resistance of the deserved load has been modified in such a way that the output voltage is maintained constant (at 400 V) regardless the system output power or irradiation level incident to the solar array. Also, a voltage and a current sensor are used for measuring the voltage across the PV modules terminals (V_{pv}) and the current through the panels (I_{pv}). The Perturb and Observe control law algorithm has been implemented using Altair Embed RCP environment and Texas Instruments C2000—LAUNCHXL—F28069M DSP evaluation board. The parameters of the MPPT algorithm used in Altair Embed are sampling time, sT = 0.5 s and perturbation step size, Delta_D = 0.01.

Table 1 presents the power stage specifications and components used for practical implementation of the IDBIC converter.

Fig. 6 Laboratory test setup

Table 1 Power stage components and specifications

Parameters	Values
Input Voltage—V_{in}	50–150 V
Output Voltage—V_{out}	400 V
Switching frequency—f_s	100 kHz
Max. Output Power—P_{out}	1200 W
Switches (T_1-T_2)	SCT3120 AL
Switches (T_3-T_4)	IRFP260PBF
Diodes (D_1, D_2)	C3D02060
Inductors (L_1, L_2)	150 μH, RL = 21 mΩ
Capacitors (C_1, C_2)	150 μF

Table 2 Maximum power of the PV array at different values of solar irradiation

Irradiation [W/m^2]	500	600	800	1000
Maximum power [W]	449.8	540.7	720.7	898.9

3 Results and Discussion

3.1 Simulation Results

Multiple simulations of the power stage operation were performed and recorded for the cases in which the profile of solar irradiation (I_r) on the PV array had the values shown in Table 2. The temperature of the PV modules was set at 25 °C.

Figure 7a represents the waveforms of the PV array power (P_{pv}), the voltage at the terminals of the panel group (V_{pv}), current through photovoltaic panels (I_{pv}) and the duty cycle of the power stage (D). In Fig. 7b the inductor currents (I_{L1}, I_{L2}) and the output voltage (V_{out}) and current (I_{out}) waveforms are presented. The simulation time intervals in correlation with the values of solar irradiations are shown in Table 3.

3.2 Experimental Results

The experimental measurements were performed in the same condition presented in the simulation results section. Figure 8 shows the oscilloscope representation for the converter inductor currents (I_{L1}, I_{L2}) and for the input and output currents and voltages waveforms at solar irradiation of 1000 W/m^2. In Fig. 9 the experimental results of the practical measurements, recorded in Altair Embed environment are presented. The waveforms of the power, voltage and current of the solar panels array are presented in correlation with the duty cycle (D) of the converter.

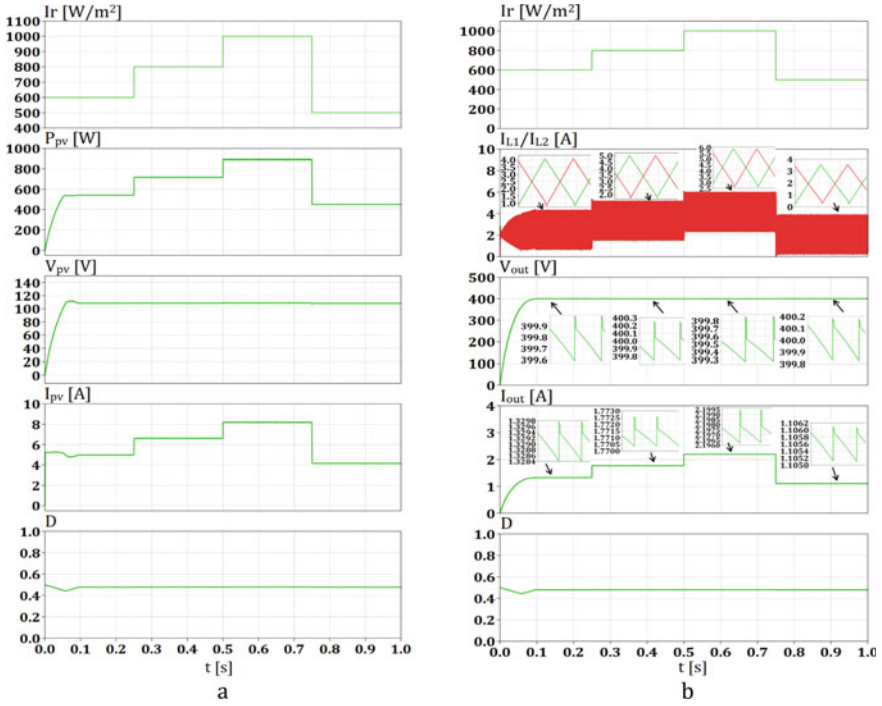


Fig. 7 Simulation results of the proposed MPPT system

Table 3 Solar irradiation levels at different simulation time intervals

Simulation time [s]	0–0.25	0.25–0.5	0.5–0.75	0.75–1
Irradiation [W/m^2]	600	800	1000	500

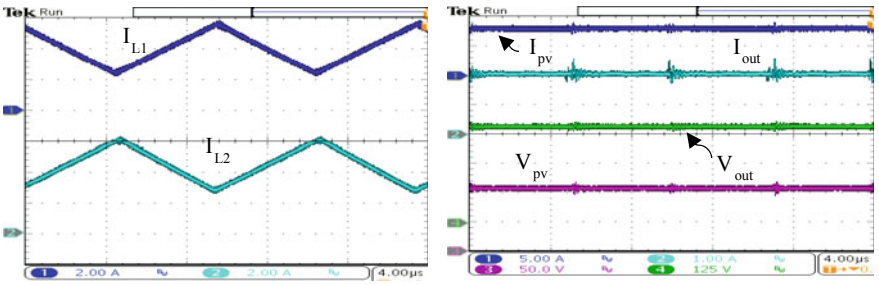


Fig. 8 IDBIC Oscilloscope resulting waveforms for $I_r = 1000$ W/m^2

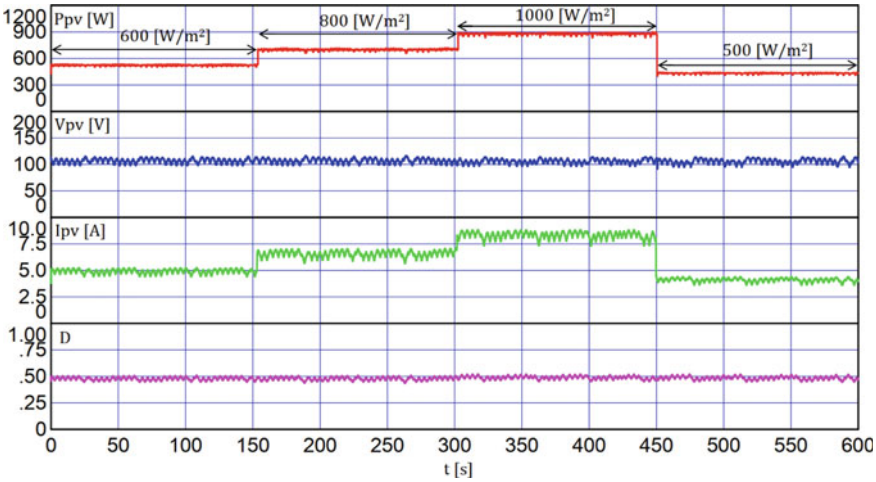


Fig. 9 Experimental results of the proposed MPPT system

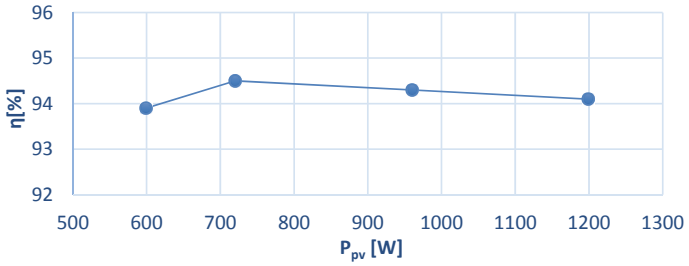


Fig. 10 Power stage efficiency versus PV array power

Using the Tektronix PA3000 power analyzer the IDBIC converter efficiency measurements was carried out for different values of solar irradiation. In Fig. 10 the relationship between the efficiency of the converter power stage and PV array power is presented.

4 Conclusion

This work emphasizes an analysis of an interleaved boost converter with three-level output working in a renewable energy application. A maximum power point tracking control law algorithm was implemented in order to extract the maximum power from the three PV modules connected is series. The power stage operation mode was analyzed by co-simulation and Real-Time Rapid Control Prototyping experimental measurements. The currents and voltage waveforms were taken for different values

of solar irradiation across three PV modules connected in series. Also, the efficiency of the power stage was determined during the Real-Time simulation, at different values of solar irradiation.

Acknowledgements This research was funded by EUROPEAN REGIONAL DEVELOPMENT FUND through the COMPETITIVENESS OPERATIONAL PROGRAM 2014–2020 Romania, grant number 16/01.09.2016, project title "High power density and high efficiency micro-inverters for renewable energy sources—MICROINV".

References

1. Çelik Ö, Teke A, Tan A (2018) Overview of micro-inverters as a challenging technology in photovoltaic applications. Renew Sustain Energy Rev 82:3191–3206
2. Luo Y, He J, Liu H, Wu L (2014) Application of the distributed generation, micro and smart power grid in the urban planning. In: The 4th annual IEEE international conference on cyber technology in automation, control and intelligent, Hong Kong, June, pp 634–637
3. Jäger-Waldau A (2019) European commission, and joint research centre. PV status report 2019 [Online]. Available. https://op.europa.eu/publication/manifestation_identifier/PUB_KJNA29938ENN. Accessed 28 Dec 2021
4. Nema P, Nema RK, Rangnekar S (2009) A current and future state of art development of hybrid energy system using wind and PV-solar: a review. Renew Sustain Energy Rev 13(8):2096–2103
5. Garg VK, Sharma S (2018) Overview on microgrid system. In: 2018 Fifth international conference on parallel, distributed and grid computing (PDGC), Solan Himachal Pradesh, India, Dec 2018, pp 694–699
6. Forouzesh M, Siwakoti YP, Gorji SA, Blaabjerg F, Lehman B (2017) Step-up DC–DC converters: a comprehensive review of voltage-boosting techniques, topologies, and applications. IEEE Trans Power Electron 32(12):9143–9178
7. Krzywinski G (2015) Integrating storage and renewable energy sources into a DC microgrid using high gain DC DC boost converters. In: 2015 IEEE First international conference on DC microgrids (ICDCM), Atlanta, GA, USA, June 2015, pp 251–256
8. Suciu VM, Salcu SI, Pacuraru AM, Pintilie LN, Szekely NC, Teodosescu PD (2021) Independent double-boost interleaved converter with three-level output. Appl Sci 11(13):5993
9. Teodosescu PD, Suciu VM, Szekely NC, Bojan AM, Zolth M (2019) Interleaved buck-boost voltage converter. OSIM Patent Application A2019 00915, 19 Dec 2019
10. Angadi S, Yaragatti UR, Suresh Y, Raju AB (2020) Implementation of perturbation-based MPPT technique using model-based design. In: Mallick PK, Meher P, Majumder A, Das SK (eds) Electronic systems and intelligent computing, vol 686. Springer, Singapore, pp 155–166
11. Kamran M, Mudassar M, Fazal MR, Asghar MU, Bilal M, Asghar R (2020) Implementation of improved perturb & observe MPPT technique with confined search space for standalone photovoltaic system. J King Saud Univ Eng Sci 32(7):432–441

Predictive Analytics and Intelligent Decision Support Systems in Supply Chain Risk Management—Research Directions for Future Studies

Patrick Brandtner

Abstract Today's supply chains (SC) are immersed in extremely dynamic environments, and supply chain management (SCM) has to deal with a multitude of risks. The domain of supply chain risk management (SCRM) has emerged, providing approaches on how to cope with risks in SC. However, due to increased complexity, volatility, and uncertainty, the number of risks in global SC has increased significantly. Harnessing the power of predictive analytics (PA), implemented in intelligent decision support systems (IDSS), offers huge potential in SCRM. However, research at the intersection of the domains of SCRM, PA, and IDSS is still in its infancy, and several research gaps have yet to be addressed. The paper elaborates on these research gaps by means of a systematic literature review. The results include a set of seven research questions and proposed research directions for future studies. Future research is presented with a plethora of starting points, which originate from the business perspective (i.e., the SCRM domain), the data-driven (i.e., the PA domain) as well as an IT-system perspective (i.e., the IDSS domain).

Keywords Supply chain management · Predictive analytics · Supply chain risk management · Decision support

1 Introduction

Logistics and supply chain management (LSCM) represent the backbone of our economy. Regardless of sectors or industries, supply chains (SC) are the preferred form of value creation in manufacturing and distributing products and services and ultimately in delivering value to customers. The evolution of SC and supply chain management (SCM) can hence be seen as a natural adoption of companies to nowadays high-competition business environments [1]. The notion of SCM goes back to Keith Oliver, who defined it in 1982 as "the process of planning, implementing, and controlling the operations of the supply chain with the purpose to satisfy customer

P. Brandtner (✉)
University of Applied Sciences Upper Austria, Steyr, Austria
e-mail: patrick.brandtner@fh-steyr.at

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_50

requirements as efficiently as possible" [2]. Considering the economic backgrounds and market dynamics given in the early 1980s and comparing it with current market and economic developments, it becomes obvious that modern SC is very different from SC of even just a few years ago [3].

High degrees of uncertainty cause shorter life cycles of nodes (entities as, e.g., suppliers, customer) and edges (relations as, e.g., "delivers" or "produces") in interorganizational SC. Long proven SCM tactics, SC structures, and relations are no longer stable and are subject to disruptive change and network dynamics [4].

Traditional logistics and SCM tactics will not be able to tackle the challenges and to deal with the large array of uncertainties in current and future SC, and the need for anticipatory, data-driven, and intelligent decision support for supply chain risk management (SCRM) is recognized in industry and academia [1, 3, 5, 6]. Particularly, high potential for uncertainty reduction and intelligent decision support systems (IDSS) in SCRM is attributed to artificial intelligence (AI) and more precisely to machine learning (ML)-based predictive analytics (PA) [7–9]. However, the potential of PA and ML has not yet been fully explored in SCRM, application studies are missing [3], and prior research mostly focused on technological aspects rather than on how decision intelligence is actually built and operationalizes in SCM [9].

The aim of this paper is to provide an overview of current research gaps at the intersection of supply chain risk management and predictive analytics-based intelligent decision support systems. More precisely, we identify three main areas of gaps in existing work and provide a list of seven research questions as potential opportunities for future research studies. To this end, we apply a systematic literature review in accordance with the systematic literature review process by Webster and Watson [10]. Literature reviews allow for developing a solid foundation and framework for subsequent research activities and provide an appropriate approach to address the aims of this paper. The main research questions addressed in the paper are defined as follows:

- Which research gaps can be observed in current literature at the intersection of SCRM, PA, and IDSS?
- What are possible research questions directions for future research studies at the intersection of SCRM, PA, and IDSS?

The remainder of the paper is structured as follows: section two provides an overview of the background and theoretical concepts (SCRM, PA, and IDSS). Section three elaborates on the research gaps at the intersection of these concepts, followed by section four, which proposes research questions that evolve out of these gaps. Finally, section five provides a discussion of results and an outlook on future research.

2 Background

The background of the three focal concepts of the paper, i.e., SCRM, PA, and IDSS is presented in the subsequent Sects. 2.1, 2.2, and 2.3.

2.1 Supply Chain Risk Management

SC risks occur in decision-making situations in SCM in which a decision-maker in an organization (i) lacks information and/or understanding of the VN and its environment, (ii) lacks the information processing capabilities needed to form decisions, (iii) is indistinct about the actual decision objectives, (iv) is unable to accurately predict possible outcomes of actions on VN behavior, or (v) lacks effective actions at all [11]. As a response to this, the domain of supply chain risk management (SCRM) has emerged, providing approaches on how to cope with risk in SCM [12]. Based on van der Vorst and Beulens (2002), the aim of SCRM should be to reduce risk, i.e., to reduce (i) lack of information about the SC, (ii) lack of information processing capabilities, (iii) lack of decision objectives, (iv) lack of prediction accuracy, and (v) lack of options for action.

The relevance of appropriately reducing and dealing with SC risks has grown steadily over time, mainly due to more complex, volatile, hardly predictable, and highly competitive markets [13, 14]. Consequently, numerous research studies have already been conducted in this research field. To cope with SC risks, a variety of approaches has been proposed by research. One approach offering huge potential in this context is predictive analytics based on machine learning models [3, 7, 15–18]. However, regarding the application of PA in SCRM, research is still in its infancy [19] and new approaches and additional research is needed [5, 7].

2.2 Predictive Analytics

Predictive analytics (PA) is one of four categories of data analytics (DA), the other three being descriptive analytics, diagnostic analytics, and prescriptive analytics. PA refers to the type of data analytics that makes predictions about yet unknown, hence uncertain, events, and risks. PA typically applies advanced analytics as AI and ML-based models. By examining past data trends and patterns in the data, PA seeks to discover the causes of events as well as to predict possible future events and risks, respectively, to fill in data or information that does not yet exist. PA makes predictions based on certain likelihoods of a risk or event, however, 100% prediction accuracy cannot be reached by PA [6]. Against the background of SCRM, PA can reduce SC risks on two levels: first, it answers the question of what will most likely happen in future SC based on patterns learned from past data. It hence generates predictive scores (probability values) for uncertain events and risks. However, PA can be applied to any type of unknown event, whether it be in the past, present, or future [20]. An example of future uncertain events in SCM includes, e.g., the probability of a delayed shipment, probable changes in customer demand, or predicted out-of-stock situations in a warehouse. PA hence captures relationships between explanatory variables (e.g., purchasing patterns of a customer) and predicted variables (e.g., the actual customer demand) and uses these relationships to predict the uncertain outcome, respectively,

the risk. The second level PA reduces SC risks are by assessing the impact of an uncertainty based on the learned patterns, respectively, relationship in the data. On example in the retail context could be to predict the lost sales (impact) based on a certain likelihood of out-of-stock (predicted variable) on the point of sale which is calculated by comparing store ordering behavior with customer demand behavior (explanatory variables) [16, 21]. Another example could be the prediction of customer churn (impact) based on observed customer behavior patterns (predicted variable) at the point of sale due to COVID-19 measures of stores (explanatory variables) [22, 23]. In summary, there is huge potential of applying PA in future, intelligent, and digital SCM [24] and specifically in the context of VN uncertainty reduction. However, research in this regard is still in a nascent stage [3, 5, 7].

2.3 Intelligent Decision Support Systems

Modern SC is immersed in extremely dynamic contexts and is surrounded by a multitude of risks—ranging from demand, supply, process, and control risks to environmental risks. One approach that has evolved out of the need to reduce risk and support decision-making with predictive intelligence based on AI, ML, and PA is the concept of intelligent decision support systems (IDSS). IDSS can be defined as decision support systems that utilize AI-based techniques to provide high-value information to a decision-maker [25]. In the context of SCRM, the potential for IDSS is huge, and such systems could fill a gap in SCM by processing large amounts of data to provide predictive statements regarding the likelihood and impact of SC risks and supporting decision-makers in an efficient and targeted way. IDSS based on data mining (DM) and PA are much more suited to fulfill the requirements of SCRM. By focusing on identifying patterns in data related to the different types of uncertainty, such IDSS apply a broad variety of PA algorithms: classification, pattern recognition, clustering, and association rule mining are just a few examples in this context. IDSS is able to solve difficult decision problems which require involving large amounts of distributed data [25]. In conclusion, IDSS and PA offer huge potential in SCRM, still, research in this context is limited, and the operationalization of IDSS in SCRM practice is challenging [26, 27].

3 Research Gaps at the Intersection of SCRM, PA, and IDSS

Based on the extensive literature analysis conducted in accordance with [10], we identify three main areas of research gaps at the intersection of the three focal theoretical concepts of this paper (i.e., SCRM, PA, and IDSS): First, we observe a lack of research in the SCRM domain that specifically focuses on PA and IDSS. The first

Table 1 Overview of research gaps and supporting references in RG 1—SCRM

Summary of research gap	Supporting references
RG 1.1: Research on how to define requirements for PA from a SCRM perspective is missing	PA is a largely underrepresented field within SCRM [28]
	AI techniques have received little attention in SCRM [29]
	Predictive AI capabilities in SCRM are in their infancy [19]
	Data-driven predictive ML is rarely examined in SCRM [30]
	Theory and applications of PA in SCRM are missing [26]
RG 1.2: Research on how to apply and implement PA in SCRM is missing	SCRM needs to incorporate PA and AI into its scope [31]
	Converting SCRM into PA problems is a challenge [17]
	Practitioners need to be involved in ML and PA for SCRM [30]
	Practitioner frameworks for PA-based SCRM are missing [19]

research gap area is hence summarized as missing research in SCRM domain (RG 1). The second area of research gaps lies in the domain of predictive analytics. More precisely, there is a lack of research on how to apply and implement PA in SCRM. The second research gap area is hence summarized as missing research in PA domain (RG 2). The third focal area of research gaps originates from the IDSS domain. We observe a lack of research on how to implement IDSS in SCRM practice. The third research gap area is summarized as missing research in IDSS domain (RG 3).

The subsequent Tables 1, 2, and 3 provide an overview of the identified research gaps in each of these areas (RG 1, RG 2, and RG 3) and state the relevant sources.

4 Research Questions for Future Research Studies

Based on the identified research gaps, we propose several research questions for future studies at the intersection of SCRM, PA, and IDSS. From an SCRM perspective, the summarized research gaps include

- RG 1.1. Application domain knowledge in SCRM to specify and define concrete requirements for PA solutions is missing.
- RG 1.1. The limited initial studies available have only explored few PA techniques and have not yet exploited newer techniques and recent advances in established ML techniques in the context of SCRM.

Table 2 Overview of research gaps and supporting references in RG 2—PA

Summary of research gap	Supporting references
RG 2.1: Research on the full technical potential and on the concrete technical specifications of PA in SCRM is low	Only few studies apply PA in SCM and even less in SCRM [3]
	AI techniques have received little attention in SCRM [29]
	Literature is scarce for SCRM on a manufacturer level [7]
	Hybrid frameworks combining PA and SCRM are needed [19]
RG 2.2: There is a knowledge gap among practitioners regarding basics of PA and their potential in SCRM	Managers lack understanding of PA in SC risk reduction [7]
	There have to be models for specific types of SC risks [32]
	There is a lack of PA understanding and skills in SC [19]
	Integrating SCM expertise into PA-based SCRM is crucial [4]

Table 3 Overview of research gaps and supporting references in RG 3—IDSS

Summary of research gap	Supporting references
RG 3.1: There is a lack of research regarding the implementation of PA capabilities in IDSS in SCM	Inclusion of PA into SCM decision-making is needed [33]
	Future work should assist SCRM practice in PA selection [3]
	IDSS in SCM have only been partially explored [34]
RG 3.2: There is a lack of research on synergetic co-existence of artificial and human intelligence in SCM	There is a need for explainable PA in SCRM [27]
	The impact of AI-IDSS on humans in SCM is unclear[3]
	New research in data-driven SCRM is needed [26]
RG 3.3: Research on how to implement and operationalize IDSS in SCM is low	Decision-making frameworks are missing in SCRM [19]
	Implementation frameworks of PA in SCRM are missing [3]
	Prior research lacks how SCM builds IDSS capabilities [9]

- RG 1.2. The application domain knowledge in SCRM has to be enriched by concrete use cases and specific application benefits of PA.
- RG 1.2. Only few studies provide implementation frameworks for PA in SCRM. The ones that do have not or only partially proceeded to implementing, applying, and evaluating their approach and are limited.

Research questions evolving out of these research gaps in the domain of SCRM are

- RQ 1 What are the specific requirements for PA from a SCRM perspective?
- RQ 2 How can PA algorithms be applied and implemented in SCRM use cases?

From a PA perspective, the following points summarize the main research gaps against the background of SCRM:

- RG 2.1. Technical domain knowledge in PA is missing in research, hindering the analysis of the full potential of PA in SCRM.
- RG 2.1. Especially in the context of demand and supply risk reduction in SC reality, research need is seen in academia.
- RG 2.1. Hybrid and parallel PA coupled with consensus seeking algorithms represent innovative approaches which are yet underexplored in SCRM.
- RG 2.2. Missing technical domain knowledge on a basic level on practitioners' sides hinders fast adoption of PA in SCRM and SCM in general.
- RG 2.2. SCM and SCRM expertise must be enriched with basic PA knowledge in practice to speed up its implementation in SC.

Building on these summarized research gaps, the following research questions for future studies can be defined in the PA domain:

- RQ 3. What are the technical capabilities and specifications of PA to fulfill the specific requirements in SCRM and how can its potential be objectively demonstrated in selected use cases?
- RQ 4. How can basic PA domain knowledge and successful PA use cases in SCRM be communicated to practice to speed up PA implementation in SCRM?

In the context of IDSS, several research gaps were identified against the background of SCM, SCRM, and PA. The following list provides a summary of these:

- RG 3.1. Existing decision support systems in SCRM lack PA capabilities.
- RG 3.1. It is yet to understand where, when, and how to use PA to support SCM objectives reduce SC risks.
- RG 3.2. Existing applications of PA in SCRM lack interpretability and a need to move toward explainable AI-based PA are given.
- RG 3.2. The impact of AI on SCM decision-makers and different degrees of combinations of AI and human intelligence is yet to be explored.
- RG 3.3. Technical domain knowledge on how organizations build IDSS capabilities in a SCM context is low.
- RG 3.3. Frameworks on how to implement and operationalize IDSS in SCM in order reduce VN uncertainties are missing.

The following research questions can be derived for the IDSS domain:

- RQ 5. How can IDSS systems be developed further to include the required PA capabilities for SCRM?
- RQ 6. How do AI generated results have to be provided to decision-makers to be understood and used in SCM and what is the impact of PA on SCM decision-makers in SC risk reduction?
- RQ 7. How can PA-based IDSS capabilities be built up in organizations and how does a framework for their implementation in SCRM have to be designed?

5 Discussion of Results and Outlook

The derived research questions represent potential starting points for future research studies at the intersection of SCRM, PA, and IDSS. In total, we have identified seven main starting points for future work, which all represent valuable contributions to research and address urgent research gaps. There is huge potential for generating new knowledge and contributing to addressing the identified research gaps and questions.

For RQ 1, the task for future studies could include the derivation of SCRM requirements in organizational practice. These requirements should be prepared in the form of concrete PA and IDSS requirements. For RQ 2, exemplary research tasks may include the analysis of SC structures, SC risk sources, and the determination of SCM objectives and success criteria. The results could be documented in the form of potential use cases for PA application in the context of SCRM. For RQ 3, the selection, application, and evaluation of PA algorithms based on actual requirements and use cases of SCRM are important exemplary research tasks. Their results may include demonstrated technical capabilities of PA in SCRM. For RQ 4, we consider a practitioner-oriented preparation and documentation of PA use cases in SCRM as important activities to foster PA application in managing SC risks. The results may be presented in the form of white papers and presented at scientific and practitioner conferences and outlets or in networks of academia and industry. The main tasks for addressing RQ 5 may comprise the elaboration of requirements for IDSS in SCRM practice. More precisely, the results could provide the basis for how to further develop existing and implement new intelligent decision support features in SCM practice. For RQ 6, we would suggest conducting an analysis and depiction of decision routines in SC risk reduction processes in practice and depict the results in the form of decision-making structures and IDSS implementation requirements in SCRM. Finally, for RQ 7, exemplary research tasks may include the development of implementation procedures for IDSS in SCM and SCRM based on analyzed SCRM use cases and PA-requirements for SCRM.

Future research is presented with a plethora of starting points at the intersection of SCRM, PA, and IDSS. These starting points may originate from a business perspective (i.e., the SCRM domain), as well as from a data-driven (i.e., the PA domain) or an IT-system perspective (i.e., the IDSS domain). Building on the identified research

gaps and questions, we have submitted a proposal for a large-scale, application-oriented research project with three industry partners. First use case drafts for each partner organization have already been developed.

References

1. Framinan JM (ed) (2022) Modelling supply chain dynamics. Springer International Publishing, Cham
2. Oliver RK, Webber MD (1982) Supply-chain management: logistics catches up with strategy. Outlook 5(1):42–47
3. Riahi Y, Saikouk T, Gunasekaran A, Badraoui I (2021) Artificial intelligence applications in supply chain: a descriptive bibliometric analysis and future research directions. Expert Syst Appl 173:114702. https://doi.org/10.1016/j.eswa.2021.114702
4. Wehberg GG (2021) Digital supply chains: key facilitator to industry 4.0 and new business models, leveraging S4 HANA and beyond. Routledge, Abingdon Oxon, New York NY
5. Stefanovic N (2021) Big data analytics in supply chain management. In: Wang J, Khosrow-Pour MDBA (eds) Advances in logistics, operations, and management science, encyclopedia of organizational knowledge, administration, and technology. IGI Global, pp 2443–2457
6. Darvazeh SS, Vanani IR, Musolu FM (2020) Big data analytics and its applications in supply chain management. New Trends Use Artif Intell Ind 4.0 p 175
7. Gonçalves JN, Cortez P, Carvalho MS, Frazão NM (2021) A multivariate approach for multi-step demand forecasting in assembly industries: empirical evidence from an automotive supply chain. Decis Support Syst 142:113452. https://doi.org/10.1016/j.dss.2020.113452
8. de Krom B (2021) Supplier disruption prediction using machine learning in production environments. Delft University of Technology, Delft, PhD-Thesis
9. Fadler M, Legner C (2020) Building business intelligence & analytics capabilities-a work system perspective. In: ICIS
10. Webster J, Watson RT (2002) Analyzing the past to prepare for the future: writing a literature review. MIS Q 26(2):xiii–xxiii. [Online]. Available: https://www.jstor.org/stable/4132319
11. van der Vorst JG, Beulens AJ (2002) Identifying sources of uncertainty to generate supply chain redesign strategies. Int J Phys Dist Log Manage 32(6):409–430. https://doi.org/10.1108/09600030210437951
12. Jüttner U, Peck H, Christopher M (2003) Supply chain risk management: outlining an agenda for future research. Int J Log Res Appl 6(4):197–210. https://doi.org/10.1080/1367556031001627016
13. Durugbo CM, Al-Balushi Z, Anouze A, Amoudi O (2020) Critical indices and model of uncertainty perception for regional supply chains: insights from a Delphi-based study. Supply Chain Manage: An Int J 25(5):549–564. https://doi.org/10.1108/SCM-10-2019-0373
14. Sato Y, Tse YK, Tan KH (2020) Managers' risk perception of supply chain uncertainties. Ind Manag Data Syst 120(9):1617–1634. https://doi.org/10.1108/IMDS-01-2020-0049
15. Brandtner P (2020) Requirements for value network foresight-supply chain uncertainty reduction. In: ISPIM conference proceedings, pp 1–12
16. Brandtner P, Darbanian F, Falatouri TN, Udokwu C (eds) (2021) Dimensions of data analytics in supply chain management: objectives, indicators and data questions
17. Kara ME, Fırat SÜO, Ghadge A (2020) A data mining-based framework for supply chain risk management. Comput Ind Eng 139:105570
18. Malviya L, Shreemali J, Ojha R, Chakrabarti P, Poddar S (2021) Transformer prediction in the supply chain using machine learning. Mater Today: Proc. https://doi.org/10.1016/j.matpr.2020.12.625

19. Baryannis G, Validi S, Dani S, Antoniou G (2019) Supply chain risk management and artificial intelligence: state of the art and future research directions. Int J Prod Res 57(7):2179–2202. https://doi.org/10.1080/00207543.2018.1530476
20. Ongsulee P, Chotchaung V, Bamrungsi E, Rodcheewit T (2018) Big data, predictive analytics and machine learning. In: Big data, predictive analytics and machine learning, pp 1–6
21. Brandtner P, Darbanian F, Falatouri T, Udokwu C (2021) Improving demand prediction and reducing out-of-stock—application of advanced data analytics in retail supply chains. In: Proceedings of the 8th international physical internet conference, p 9
22. Brandtner P, Darbanian F, Falatouri T, Udokwu C (2021) Impact of COVID-19 on the customer end of retail supply chains: a big data analysis of consumer satisfaction. Sustainability 13(3):1464
23. Udokwu C, Darbanian F, Falatouri TN, Brandtner P (2020) Evaluating technique for capturing customer satisfaction data in retail supply chain. In: 2020 The 4th international conference on e-commerce, e-business and e-government. Arenthon France, pp 89–95
24. Jandhyala R (2021) Intelligent digital supply chains. In: Srinivas S, Rajendran S, Ziegler H (eds) International series in operations research & management science, supply chain management in manufacturing and service systems. Springer International Publishing, Cham, pp 19–64
25. Phillips-Wren G (2013) Intelligent decision support systems. In: Doumpos M, Grigoroudis E (eds) Multicriteria decision Aid and artificial intelligence. John Wiley & Sons, Ltd, Chichester, UK, pp 25–44
26. Ivanov D, Dolgui A (2021) A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. Prod Plan Control 32(9):775–788. https://doi.org/10.1080/09537287.2020.1768450
27. Nimmy SF, Hussain OK, Chakrabortty RK, Hussain FK, Saberi M (2021) Explainability in supply chain operational risk management: a systematic literature review. Knowl-Based Syst 107587. https://doi.org/10.1016/j.knosys.2021.107587
28. Stampe L, Hellingrath B (2021) Risk indicators and data analytics in supply chain risk monitoring. In: Buscher U, Lasch R, Schönberger J (eds) Lecture notes in logistics, logistics management. Springer International Publishing, Cham, pp 246–263
29. Tirkolaee EB, Sadeghi S, Mooseloo FM, Vandchali HR, Aeini S (2021) Application of machine learning in supply chain management: a comprehensive overview of the main areas. Math Probl Eng 2021:1–14. https://doi.org/10.1155/2021/1476043
30. Engelking B, Buchholz W, Köhne F (2020) Design principles for the application of machine learning in supply chain risk management: an action design research approach. In: Bode C, Bogaschewsky R, Eßig M, Lasch R, Stölzle W (eds) Advanced studies in supply management, supply management research. Springer Fachmedien Wiesbaden, Wiesbaden, pp 137–162
31. Fagundes MVC, Teles EO, Vieira de Melo SA, Freires FGM (2020) Decision-making models and support systems for supply chain risk: literature mapping and future research agenda. Eur Res Manag Bus Econ 26(2):63–70. https://doi.org/10.1016/j.iedeen.2020.02.001
32. Sarker S (2019) The paradox of risk management: a supply management practice perspective. In: Zsidisin GA, Henke M (eds) Revisiting supply chain risk. Springer International Publishing, Cham, pp 421–437
33. Araz OM, Choi T-M, Olson DL, Salman FS (2020) Role of analytics for operational risk management in the era of big data. Decis Sci 51(6):1320–1346. https://doi.org/10.1111/deci.12451
34. Baryannis G, Dani S, Antoniou G (2019) Predicting supply chain risks using machine learning: the trade-off between performance and interpretability. Futur Gener Comput Syst 101:993–1004. https://doi.org/10.1016/j.future.2019.07.059

Impact of COVID-19 on the Belfast Bike Sharing Scheme

Lucy Doyle, Aleksandar Novakovic, Adele H. Marshall, and Darren Cheung

Abstract The COVID-19 pandemic brought the connectivity and mobility of Northern Irish residents to a halt forcing the bike hiring service Belfast Bikes to cease operations. After reintroducing the service in Summer 2020, the usage of the bike hiring service was severely impacted. This paper investigates the usage of Belfast Bikes pre-pandemic and during the pandemic. Research includes how consumer trends for Belfast Bikes customers changed due to the pandemic and predictive modelling is used to predict whether customers hire a Belfast Bike for direct travel from stations or for an indirect, leisurely trip. We will conclude how the connectivity and mobility of Belfast residents have changed due to the pandemic and provide recommendations for Belfast Bikes and Belfast City to recover after the pandemic.

Keywords COVID-19 · Trip classification · Belfast Bikes · Bike sharing prediction

1 Introduction

The first recorded public bike sharing scheme was introduced in Amsterdam, Netherlands, in 1965 as WhiteBikes. The service was introduced by Luud Schimmelpennink who encouraged the public to paint old bicycles white and leave the bikes anywhere in Amsterdam for public use. Since then, bike sharing has spread globally, with nearly

L. Doyle (✉) · A. Novakovic · A. H. Marshall
School of Mathematics and Physics, Queen's University Belfast, Belfast, UK
e-mail: ldoyle13@qub.ac.uk

A. Novakovic
e-mail: a.novakovic@qub.ac.uk

A. H. Marshall
e-mail: a.h.marshall@qub.ac.uk

A. Novakovic · A. H. Marshall
Faculty of Business and IT, Ontario Tech University, Oshawa, Canada

D. Cheung
Data, Discovery and Decision Science, Allstate NI, Belfast, UK
e-mail: dcheung02@qub.ac.uk

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_51

3000 schemes and 10 million rental bikes worldwide by mid-2021 [1, 2]. Public bicycle services offer an environmentally friendly alternative to travel through cities, whilst improving the general public's health. The bikes can be used by the public such as commuters travelling to work or tourists taking leisurely bike rides around a city. Modern bike hiring services have stations located across cities for the public to hire or return a bike with electronic bike locks or integrated lock systems. Payment is usually made through smartphone apps, with initially cheap rates to rent the bikes [3].

To the detriment of bike hiring services, the COVID-19 pandemic beginning 11th March 2020, brought the mobility of people worldwide to a halt, with working from home measures introduced and limited travelling allowed. A publicly shared service such as bike hiring created public uneasiness due to potential hygiene concerns. In the United Kingdom, the Prime Minister introduced strict stay at home orders in March 2020 for all residents of Northern Ireland (NI) [4]. The restrictions remained in place until mid-June 2020 when shops began to reopen [5], and people socialised outside their households [6]. Following this, restrictions in NI continued to change from June 2020 to May 2021. A short month-long lockdown was introduced in NI beginning 16th October 2020 [7], and a four-month lockdown started on 26th December 2020 [8]. This paper will investigate the impact of the tumultuous time from March 2020 to May 2021 on the bike hiring service Belfast Bikes with focus on:

- The impact the pandemic had on the usage of Belfast Bikes
- Change in consumer trends with Belfast Bikes due to COVID-19
- Predicting how customers will use the Belfast Bikes service in the future.

This paper is organised as follows; Sect. 2 provides background to related work and describes our approach. Section 3 explains the methods used to complete exploratory analysis of Belfast Bikes data, documentation of external data sources used and steps to build a predictive model. In Sect. 4, we state the findings, whilst Sect. 5 discusses how the results explain the impact of the pandemic on bike usage, how consumer habits have changed during 2020–21 and how we can predict future usage of the bike hiring service. The paper concludes with appropriate next steps for Belfast Bikes to recover from the reduced mobility during the pandemic.

2 Literature Review

Similar analysis on the impact of COVID-19 on bike sharing schemes investigate the immediate impact of the COVID-19 pandemic shortly after it was declared by the World Health Organisation on 11th March 2020. Shang et al. [9] analysed bike usage and bike distribution for bike sharing schemes in Beijing, China from January to March 2020. Texieria et al. [10] surveyed 294 residents in 2020, 86% residing in Lisbon, Portugal, to find their motivations in using bike sharing schemes during the pandemic. They found that GIRA shared bicycles had become the main transport for this sample of people after the pandemic, compared to public transport, the

preferred option in 2019. Similarly, De vos [11] analysed the potential impact of social distancing on consumer travel behaviours and concluded that public transport will be avoided whilst more active transport options, perhaps walking or cycling, rising in popularity.

Previously, we proposed a modelling approach aimed at predicting the number of bikes and racks to satisfy user demand for a certain period until the next relocation process [3]. The approach presented in this paper differs from previous research by considering the long-term impact of the COVID-19 pandemic on the usage of bike hiring services with exploratory analysis of the Belfast Bikes data pre-pandemic and for an extended period during the pandemic (from December 2019 to May 2021). This research includes analysis on the impact of a series of lockdown restrictions in NI on bike sharing in Belfast with real-time Belfast Bikes customer trip data. Changepoint detection algorithms find dates from December 2019 to May 2021 where significant changes in Belfast Bikes usage occurs. Also, we draw conclusions on the change in consumer habits for bike sharing customers after the COVID-19 pandemic. Changes in consumer trends are investigated with Google API data as we can compare trip times of customers' bike rides using Belfast Bikes to the Google predicted trip time. Comparing trip times will classify how customers use the bike service pre and during the pandemic. Finally, data modelling will be applied to the existing trips data to determine relationships between different features when customers have hired a bike. It will classify the customer as a user of the bike service for direct trips between stations in the Belfast City centre or as a user for indirect leisure trips.

3 Methods

3.1 Data Collection and Preprocessing

Study Area. Belfast Bikes is a bike sharing programme operating in Belfast, NI. The service allows the public to rent a bike between 6 am and midnight, 365 days a year [3]. There are 350 bikes available to pick up and drop off at 45 official bike stations across Belfast. Customers can choose a "pay-as-you-go" payment option, where each 30 min period is priced at £1. Alternatively, there is an annual membership or a casual 3 day membership available. The membership options offer the first 30 min per rental free with additional charges occurring after that. Due to the COVID-19 pandemic, Belfast Bikes ceased operations from early April 2020 to 3rd July 2020. In December 2020, Belfast Bikes introduced 450 new bikes with upgraded features using a smart technology frame lock, solar energy and GPS technology.

Data Processing Workflow. Figure 1 illustrates the data processing workflow which includes data collection and processing, feature selection for predictive modelling and the machine learning models used. Initial analysis explores Belfast Bikes data of

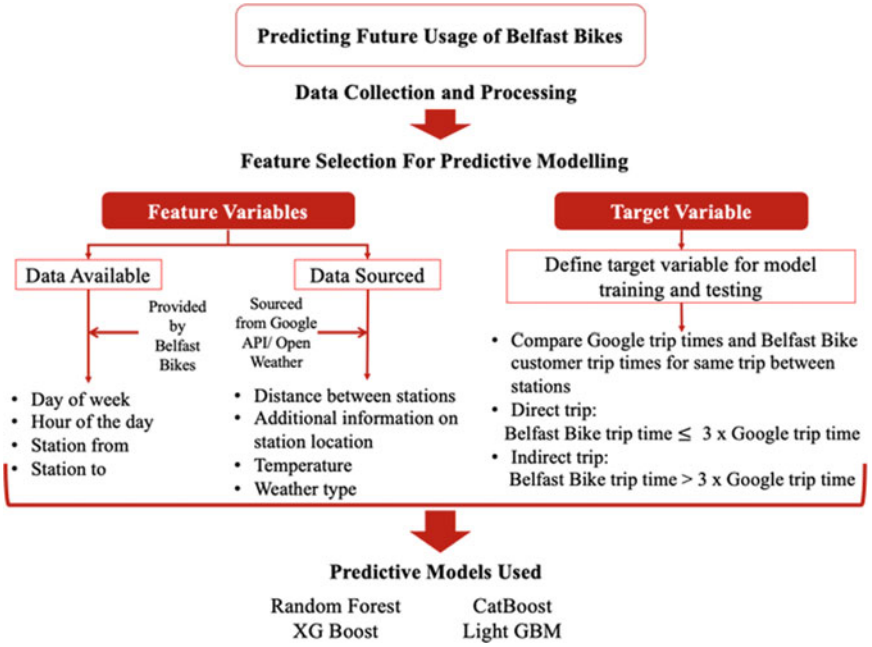


Fig. 1 Workflow diagram for feature selection and model building

bike trips to analyse Belfast Bikes usage in December 2019 to April 2021. External data is also used to identify features and target variables for modelling Belfast Bikes usage.

Bike Trips Information. Data for Belfast Bikes rides was obtained from OpenDataNI which provides live station information updated every 15 min in JSON format [12]. The cleaned data includes the date and time information it was registered, the time a journey began, the station from and time when a bike was hired and the station and time when the bike was returned.

Weather Information. Weather data was sourced from OpenWeather to investigate changes in Belfast Bikes usage due to environmental factors [13]. The temperature and weather type, such as wind, rain or clear skies, are used as features in the modelling section of the report to predict how a customer uses a Belfast Bike.

Location information. Google Maps Platform SDKs and APIs provide data on maps, routes, and places [14]. We can compare cycling travel times between stations as predicted by Google with Belfast Bikes' customers' trip times. Google Place provides information on the area surrounding Belfast Bikes stations by labelling a location as a route, intersection, point of interest, location of health, park, university or premise.

COVID-19 related data. Coronavirus related data sourced from GOV.UK includes the number of new COVID-19 cases in NI and the number of people vaccinated in NI.

3.2 Analysing COVID-19 Impact on Belfast Bikes Sharing Scheme

To identify a significant change in Belfast Bikes usage due to COVID-19, we apply changepoint analysis to the number of Belfast Bikes trips from January 2020 to early May 2021. Changepoint analysis identifies when a significant change has occurred for time-series data [15]. Offline change point detection, also known as retrospective, is applied in this report. Offline changepoint analysis is most suitable as all data is received and processed to segment when any change is detected. Online, however, is used for real-time data that is constantly updating to detect possible anomalies. We apply sliding window changepoint analysis. The sliding window algorithm calculates the difference in signal between two side by side windows. In this example, the signal is the usage of Belfast Bikes. The discrepancy between any two windows depends on a cost parameter. Once all discrepancies are calculated, a discrepancy curve is plotted, and the peaks of this curve identify when the most change in the signal occurs; these peaks are then considered the change points. The change points identified can be cross-referenced with the timeline of the COVID-19 pandemic in Northern Ireland. Alternative changepoint detection algorithms include PELT search method, binary segmentation search and dynamic programming search [15]. The sliding window algorithm was superior in identifying small changes in the number of daily bike rides for the Belfast Bikes data, whilst not producing noisy results. We implemented the changepoint analysis using the Python package "ruptures".

3.3 Predicting Belfast Bikes Usage

Development of the predictive model. Predictive modelling will classify a Belfast Bikes trip using existing trip data to predict how customers use the service. A binary classification will categorise Belfast Bikes trips into two groups:

- **Trip type one**: direct trips from station A to station B. These are labelled "direct" trips. For example, a customer using a Belfast Bike to travel to work.
- **Trip type two**: indirect trips from station A to station B. These trips are labelled "indirect" trips as they are much longer than the expected trip time from A to B. These could be leisure trips where customers go on extended bike rides.

Several machine learning tree-based algorithms were explored including Random Forest [16], XGBoost [17], CatBoost [18] and Light GBM [19].

Features. We carried out several model iterations with a different equation of classifying a trip as "direct" or "indirect" for training and testing based on a customer's trip time. The final iteration classified a trip using a multiplier cut-off value as it is more specific to the trip than an arbitrary numeric cut-off value. Therefore, a trip is classified as type one when the trip time is less than three times the predicted trip time from station to station calculated by Google Maps API. To predict a Belfast Bikes trip as "direct" or "indirect" the following features were used:

- **Trip-based**: day of the week trip occurs, hour of the day bike is hired, station bike is hired from, station bike is returned to.
- **Location data**: distance between stations, amenities at the station location.
- **Weather data**: temperature of Belfast when trip begins and weather type.

Validation. We used an 80/20 train-test split to train and test the models' classification ability and chose the best model iteration based on the AUC score. The area under the curve (AUC) score measures a model's ability to distinguish between classes, as opposed to the accuracy which measures the total correctly classified items.

4 Results

4.1 Data Insights

Trip data obtained from OpenDataNI contained 98,311 unique Belfast Bikes customer trips between midnights on 30th December 2019 to 3rd May 2021. After refining for the shortest 10% of trip times, trips <1 min, and the longest 5%, trips >188 min, 74,193 records remained for trips between 47 stations and across 625 bikes.

Figure 2 shows that little to no bike rides occur in April, May and June 2020 due to Belfast Bikes closing operations due to COVID-19. The most popular months to rent a bike were March and April 2021. Between 30th December 2019 and 1st April 2020, 19,401 Belfast Bikes hires were made. On an average weekday 273 bikes were hired, whilst on a weekend 92 bikes on average were hired. Between 30th December 2020 and 1st April 2021, 25,711 bikes were hired. For this period, on average, customers hired 301 bikes on a weekday and 270 bikes on a weekend. The raw data was filtered to remove trips less than 1 min as these trips were impossibly short to occur between two stations. The longest 5% of Belfast Bikes customer trips ranged from 188 to 45,043 min (5572 trips) and did not reflect the general consumer behaviour whilst generating noisy results when performing modelling. Therefore, the top 5% of trip times were filtered. The median time taken to travel from stations A to B for all trips recorded was 13 min. Fifty percent of trip times fell between 6 and 38 min.

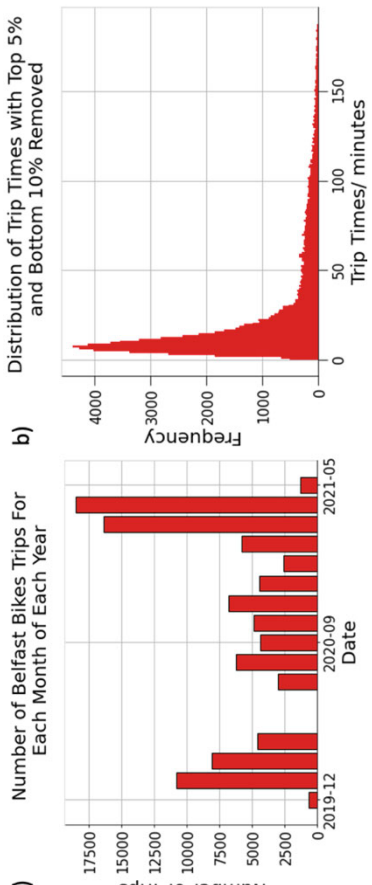


Fig. 2 Belfast Bikes data from 30th December 2019 to 3rd May 2021: **a** total monthly bike rides by customers using Belfast Bikes; **b** distribution of the duration of Belfast Bikes trips

4.2 *Analysing COVID-19 Impact on Belfast Bikes Sharing Scheme*

Sliding window changepoint detection identified five turning points for the usage of Belfast Bikes, as seen in Fig. 3. The changepoint analysis identified a significant change in Belfast Bikes usage on 13th March, 6th July and 24th September 2020, 26th February and 6th May 2021. These changepoint dates can be cross-referenced with the timeline of NI COVID-19 restrictions. The Pearson correlation between daily COVID-19 cases and the number of daily Belfast Bikes hires is −0.48. The NI vaccine roll-out began on 8th December 2020. There is a correlation of 0.82 between the number of NI people vaccinated and the number of daily Belfast Bikes hires.

The average difference between Belfast Bikes customers' trip times and Google predicted trip times is 6 min, where customers' trips are 6 min longer than Google predicted. Considering trip times for bike hires from 30th December 2019 to 1st April 2020, Belfast Bikes customers' median pre-pandemic trip time was 10 min, whilst Google predicted these trips to take 7 min on average. The Belfast Bikes trips for the

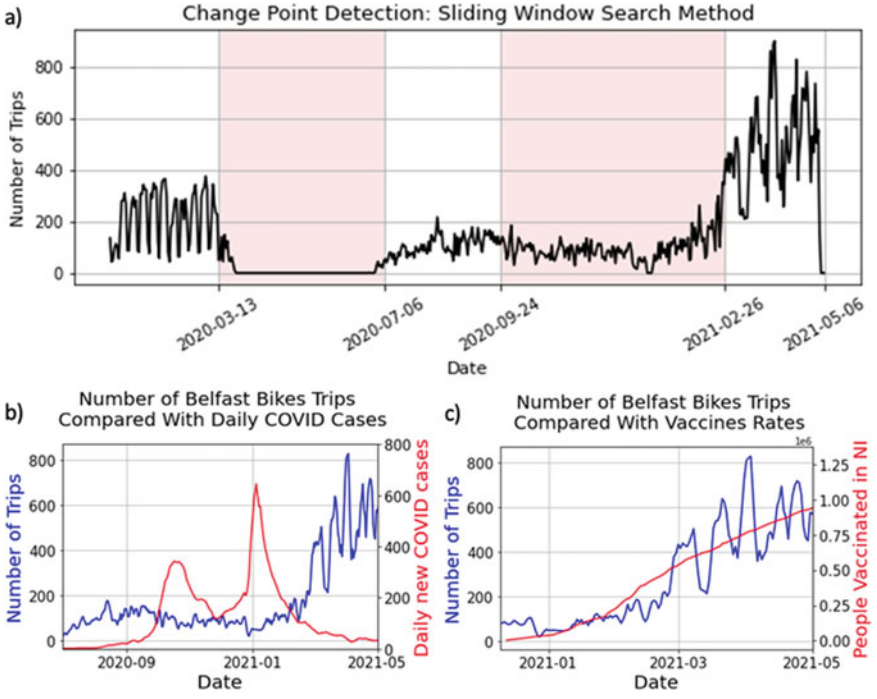


Fig. 3 Daily Belfast Bike rides for 19th December 2019 to 3rd May 2021: **a** changepoint analysis of daily Belfast Bikes trips with dates when a significant change in daily bike rides occurs; **b** daily number of Belfast Bikes trips compared with COVID-19 cases from the return of Belfast Bikes on 1st July 2020; **c** daily Belfast Bikes trips compared to the total number of NI residents with a COVID-19 vaccine, (vaccine programme began December 2020)

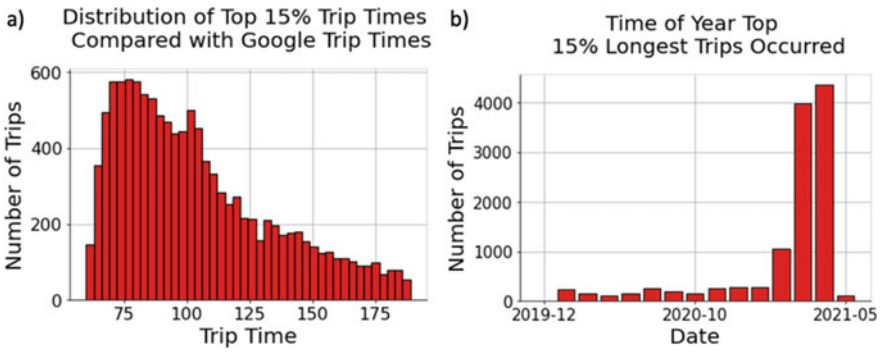


Fig. 4 The top 15% longest trips made by Belfast Bikes' customers between December 2019 and May 2021; **a** the distribution of bicycle trip times for the top 15% longest trips made by Belfast Bikes' customers; **b** the time of year when the 15% longest Belfast Bikes trips occur

same time period during the pandemic, again calculated by Google to be on average 7 min, had a median trip time of 21 min. The top 15% longest Belfast Bikes trips when compared to Google trip times were investigated. The top 15% longest trips equal 11,632 trips, and they are at least 60 min longer than Google predicted. Figure 4 shows of the 11,632 long trips, over 70% occur on or after 1st March 2021. Customers most commonly have long trips between 2 and 4 pm. Most of the trips occur on Saturdays and then uniformly across the other weekdays. The long trips are most commonly from the Queen's University/Botanic Gardens station (11.3%). The following most likely stations for a long trip will begin from are Queen's University Road and Donegall Quay, with 8.1% and 5.8% of trips leaving these stations, respectively. 10.9% of the long trips finish at Queen's University/Botanic Gardens station, closely followed by 9.5% of bikes being returned to Queen's University Road station and 6.1% returned to Titanic Belfast Met.

4.3 Predicting Belfast Bikes Usage

We classify 46,482 trips as direct (63%) and 27,711 (37%) as indirect for the training and testing of machine learning models. The distribution of the target variable is seen in Fig. 5. Table 1 shows the train and test metrics for predicting a Belfast Bikes trip as direct or indirect for four tree-based machine learning models. Considering the test data metrics, the accuracy is best for the XGBoost model but the AUC does not perform as well in distinguishing between the two classes as the Random Forest model. Therefore, the Random Forest model is selected as the best with AUC of 0.817.

Feature Importance. The random forest most predictive features are the distance between two stations, the journey start hour of the day, and the day of the week

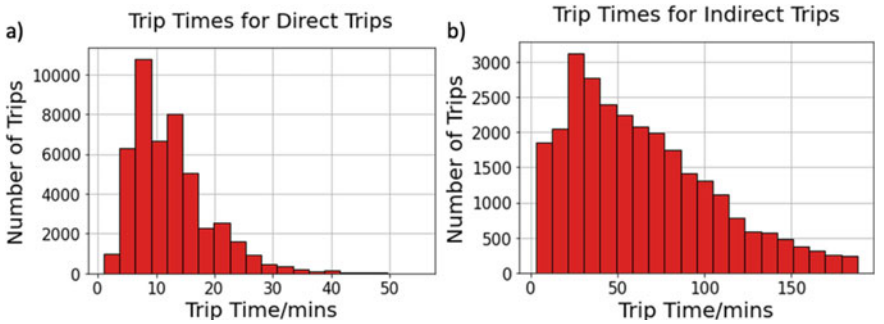


Fig. 5 The target variable for predictive modelling of a direct or indirect trip where direct is defined as a trip time less than three times the predicted trip time from station to station calculated by Google Maps API; **a** distribution of the 63% trips labelled as direct; **b** distribution of the 37% trips labelled as indirect

Table 1 Model metrics for tree-based classification models

	Training data		Test data	
	AUC	Accuracy	AUC	Accuracy
Random forest	0.989	0.990	0.817	0.841
XGBoost	0.963	0.973	0.804	0.850
CatBoost	0.988	0.987	0.810	0.847
Light GBM	0.872	0.906	0.767	0.822

a customer hires the Belfast Bikes (Fig. 6). For customers hiring Belfast Bikes for journeys between stations with a short distance, <1.1 miles, the journey is likely to be indirect whereas trips with bike stations with a distance between 1.1 and 1.7 miles

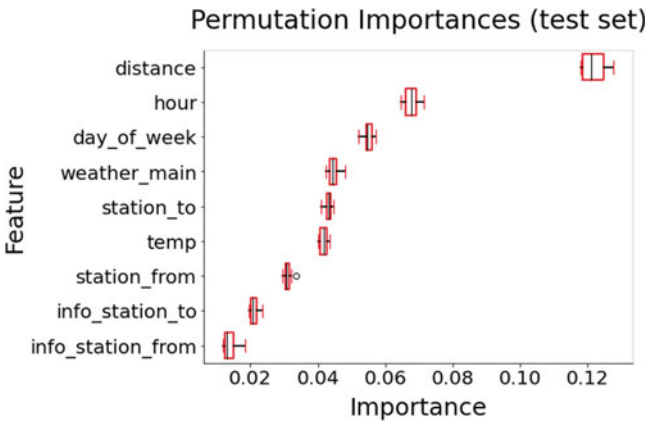


Fig. 6 Most important features when predicting a Belfast Bikes trip as a direct or indirect trip

are likely to be direct trips. Trips occurring in the morning between 8 and 10 am and late afternoon between 4 and 6 pm are likely to be direct trips. Trips occurring on weekends are likely to be indirect, whilst weekday trips are overall more likely to be direct trips. Less important features such as station information and the temperature are highly correlated to the station travelling from and to, and the main weather type, respectively. Therefore, these features do not significantly improve the model predictiveness.

5 Discussion

5.1 Impact of the Pandemic on Belfast Bikes Usage

The result of closing operations brought bike hiring to a halt for April-June 2020. Changepoint analysis identified a significant fall in Belfast Bikes usage on 13th March 2020, which was days prior to the UK nationwide lockdown and 19 days before Belfast Bikes stopped operations. The usage began to increase on 6th July 2020, almost a week after Belfast Bikes reinstated the service. Figure 3 shows reduced hiring during summer 2020 compared to bike hiring pre-pandemic in January and February 2020. Belfast Bikes could have improved usage during summer 2020 with increased marketing announcing the reinstated service and by improving consumer confidence, for example, with promises of deep cleaning of bikes. Usage significantly falls after 24th September 2020; this corresponds to drink only pubs reopening when people were welcome indoors again and the beginning of Autumn, 22nd September 2020. Due to the cold weather and relaxed restrictions, COVID-19 cases begin to increase at the end of September. The massive increase in COVID-19 cases towards the end of 2020, colder weather, and a four-week lockdown beginning 16th October 2020, perhaps discouraged Belfast Bikes hiring until 16th February 2021. After then, usage of Belfast Bikes increased with peak usage in March–April 2021 for the entire dataset.

5.2 Change in Belfast Bikes Consumer Trends Due to COVID-19

To compare Belfast Bikes usage pre-pandemic and during pandemic, due to the limited timeline of the data, comparisons can only be made between 30th December 2019–1st April 2020, and 30th December 2020–1st April 2021. Figure 2 shows direct comparisons between these two-time frames for the monthly number of trips. Pre-pandemic customers were more likely to hire Belfast Bikes on weekdays. Since Belfast Bikes' stations are in Belfast City centre, it is safe to assume the bikes are used on weekdays for customers to travel for work purposes. However, during the

end of the pandemic, the bike hiring service is equally popular on weekdays and the weekend. The increased usage of bikes on weekends implies a change in consumer trends where customers are transitioning from bike usage for work, to using the bikes for leisure.

The most popular Belfast Bikes trips include stations near education institutions and tourist attractions such as Queen's University Belfast and the Titanic Quarter. The increased popularity of these locations is despite educational institutions moving operations to work from home and tourist attractions being closed for most of April 2020 to May 2021. Therefore, a probable reason these stations are most popular is the facilities in the surrounding area. Queen's University stations are located close to the park Botanic Gardens and cafes which adapted to the pandemic by offering takeaway services. Comparing the Google predicted trip times for the most popular journeys versus the Belfast Bikes customer trip times, the times for at least 4 of the most popular journeys far exceed the average bicycle trip time calculated by Google. These long trip times imply that customers are taking bike rides for leisure instead of direct trips.

The top 15% most extended Belfast Bikes trips customers are pictured in Fig. 4. 70% of the longest trips occur on or after 1st March 2021. Most long trips occurring at this time of year shows a change in consumer trends due to the pandemic with trips during lockdown lasting anywhere between an hour and over three hours. Since these long trip times far exceed the maximum Google predicted direct trip of 21 min, we can assume these extended trips are deliberate, leisurely bike rides. Perhaps the Belfast Bikes customers meet friends whilst on their bike ride or stop at a cafe.

By 1st March 2021, NI residents are now two months into a four-month lockdown, so people may have become less tolerant of restrictions. As well, by March 2021, the COVID-19 vaccine programme is well established in NI. Therefore, boredom at home, hope for change and people feeling protected against COVID-19 through vaccination has encouraged Belfast Bikes customers to hire bikes for indirect, leisure trips. Lastly, an update to the old, often broken bikes in December 2020 meant bike stations were better populated and the new smart-lock system allowed customers to lock a bike at a location different from a station whilst hiring it. This means customers could perhaps use do other activities whilst preventing bike theft, which could extend trip times.

5.3 Predicting How Customers Will Use Belfast Bikes in the Future

Prior to the pandemic, Belfast Bikes' target customers were people aiming to quickly travel across Belfast, for example to get to work or tourists using the bikes as a method to get around the city. However, the increased usage of Belfast Bikes by local NI residents during the pandemic on weekends and around tourist hotspots, along with

increased trip times, has identified another market for Belfast Bikes to target in local NI residents who are looking for activities to get out of the house during COVID-19.

One disadvantage of Belfast Bikes customers increasingly using the bikes for longer trips is that it may be difficult to appropriately populate stations with the correct number of bikes. Longer trips keep bikes out of stations and unavailable for others. Therefore, it is beneficial for Belfast Bikes to identify trips they expect to be indirect to calculate the trip time. If Belfast Bikes expect a customer not to return a bike for an extended amount of time, then they can use other available bikes to replace them.

The predictive model described in Sect. 4.3 can predict whether a trip will be direct or indirect with customer of the stations they are renting the bike from and where they are likely to return. The most important feature in classifying a trip the distance between the two stations; therefore, a likely station to return the bike will be necessary for accurate classification. However, the hour of the day, the day of the week and the weather type are independent of where the bike is parked, therefore if the customer decides to park their bike elsewhere, Belfast Bikes can still be confident in the classification. Appropriately populating stations with bikes and marketing customers to take indirect, leisure trips will increase revenue for Belfast Bikes. Diversifying to target a wider audience of NI residents also decreases risk of for Belfast Bikes.

6 Conclusion

The stay at home orders introduced to Northern Ireland on 23rd March 2020, brought connectivity and mobility of NI residents to a halt, which is reflected in the usage of the bike hiring service, Belfast Bikes. The low number of bike hires from 1st July 2020, until the end of 2020 shows Belfast Bike's existing customer base, which mainly consisted of direct bike rides on weekdays (possibly travel for work) and tourists, were not sustainable for the service during uncertain times such as the COVID-19 pandemic. This tumultuous time gives Belfast Bikes an opportunity to re-valuate business operations and adjust for other unexpected events.

One possible change to the business could be targeting NI residents to hire Belfast Bikes for long, leisurely trips around Belfast. The leisurely trips increase sales revenue as customers are encouraged to rent the bikes for longer, and existing Belfast Bikes members will use the bike longer than the initial free 30 min. Targeting new customers may create bike supply issues, especially as long bike rides mean bikes will be out of stations for longer. Belfast Bikes can use predictive modelling such as the regression tree classifier in Sect. 4.3 to ensure that bike stations are populated enough to allow customers to rent bikes and have space to return bikes. When the model predicts a trip to be indirect, Belfast bikes can expect the bike to be returned to the intended station at least three times later than it would take to travel from station A to B. Therefore, Belfast Bikes know to fill station B if it is empty, as a bike will not be returned there for some time.

One potential limitation of our work may be that due to the unprecedented nature of COVID-19 training predictive models on the current data may not accurately reflect future Belfast Bikes usage. Therefore, in the future more recent data should be included in the training of predictive models to reflect customers bike hiring usage.

Acknowledgements The authors wish to acknowledge Allstate for proposing the challenge and for their valuable feedback throughout the project and Queen's University Belfast for funding and academic guidance and supervision.

References

1. Meddin R, DeMaio P, O'Brien O, Rabello R, Yu C, Seamon J, Benic-cio T, Han D, Mason J (2009) The Meddin bike-sharing world map. https://bikesharingworldmap.com
2. DeMaio P (2009) Bike-sharing: history, impacts, models of provision, and future. J Public Transport 12(4):41–56
3. Demidova N, Novakovic A, Marshall AH (2022) Optimizing the belfast bike sharing scheme. In: Intelligent systems and applications. IntelliSys 2021. Lecture notes in networks and systems, vol 295
4. Johnson B (2020) Prime Minister's statement on coronavirus (COVID-19), 23 March 2020. https://bit.ly/3u6aNys
5. McCormack J (2020) Coronavirus: shops in NI reopen their doors on Friday. https://bbc.in/3AEgTHF
6. GOV.UK (2020) Meeting people from outside your household. https://bit.ly/34di3Ob
7. McCormack J (2020) Schools to close and tight new hospitality rules in Northern Ireland. https://bbc.in/3G7uCaI
8. McCormack J (2020) Coronavirus: Northern Ireland begins six-week lockdown. https://bbc.in/3KLoRmO
9. Shanga W-L, Chen J, Bi H, Sui Y, Chen Y, Yu H (2021) Impacts of covid-19 pandemic on user behaviors and environmental benefits of bikesharing: a big-data analysis. Appl Energ 285
10. Teixeira JF, Silva C, Śa FME (2021) The motivations for using bike sharing during the covid-19 pandemic: insights from Lisbon. Transp Res Part F Traffic Psychol Behav 82
11. Vos JD (2020) The effect of covid-19 and subsequent social distancing on travel behaviour. Transp Res Interdisciplinary Perspec 5
12. OpenDataNI Belfast bike live data. https://bit.ly/3r5c9ay
13. OpenWeather Current forecast weather data collection. https://openweathermap.org/api
14. GoogleDevelopers Google maps platform documentation. https://bit.ly/3r6ZSTl
15. Truonga C, Oudreb L, Vayatis N (2020) Selective review of offline changepoint detection methods. Sig Process 167
16. Breiman L (2001) Random forests. Mach Learn 45:5–32
17. Chen T, Guestrin C (2016) Xgboost. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining
18. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A (2017) Catboost: unbiased boosting with categorical features. CoRR
19. Ke G, Meng Q, Finley T, Wei TW, Ma CW, Ye Q, Liu T-Y (2017) Lightgbm: a highly efficient gradient boosting decision tree. In: Advances in neural information processing systems, vol 30

Multilingual Complementation of Causality Property on Wikidata Based on GPT-3

Yuxi Jin and Shun Shiramatsu

Abstract We aim to develop an agent for understanding the distribution of public opinions and preferences; thus, this agent needs to have causality knowledge. When discussing social issues, Wikidata, a knowledge base, can provide linked data and play an important role in analyzing discussion content. However, there is a lack of causal content on Wikidata, and some content has errors. Therefore, it is necessary to automatically extract knowledge from news and add it to Wikidata. We propose a method of automatically determining causality in text and directly extracting effect from news. We collected news and used GPT-3 to infer whether a news article is causally related to the entity and further infer the effect of this entity. We also attempted to increase the reliability of extracted causality knowledge by dealing with multilingual texts.

Keywords GPT-3 · Knowledge graph · Wikidata · Causality · Information extraction

1 Introduction

People have become increasingly concerned about social problems, and many voluntary organizations are discussing such problems. Clarifying the causes of social problems and understanding their cause–effect relationships can contribute to the public debate, ideation, and public collaboration for the solution. We are aiming to develop a software agent to support such public debate, which will use Wikidata as a knowledge base for understanding public opinions and the causal structure behind the social problems. However, many Wikidata entries are missing or have incorrect causality. It depends on only special cases, not general causality, and cannot provide reliable knowledge specifically for analyzing the discussion content.

Y. Jin (✉) · S. Shiramatsu
Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Aichi, Japan
e-mail: jinyuxi@srmtlab.org

S. Shiramatsu
e-mail: siramatu@srmtlab.org

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_52

We propose a method of automatically determining causality in a text and extracting effect in news triggered by specific words. To obtain general causality, we use our method to predict a large number of news items of the same entity and group the predicted results to find the general results of high-frequency occurrences. To improve the trustworthiness of multiple extracted results, we summarize the extracted results from news in English, Japanese, and Chinese to finally obtain the general results with high trustworthiness.

The contributions of our research are as follows:

- Different from the traditional model that uses semantic features, vocabulary symbol features, and self-construction constraints to collect causality, we use the generative pretrained transformer 3 (GPT-3) [1], which is effective for ambiguous expression of causal content.
- We used three languages to avoid inferring special case effects from news and calculate the reliability scores of the results, laying the groundwork for later adding the general results to Wikidata.
- Although, our purpose is to extract "effect" of the entity from news, we believe that our method can be extended to extract other properties like "cause," "subclass," and so on. It provides reference value for automatically perfecting various properties on Wikidata.

2 Related Works

There are three main types of methods for causal extraction: pattern matching [2–4], combined pattern matching and machine learning [5], and deep learning [6–8, 10]. Compared with the traditional method of extracting causality through pattern matching using semantic features, vocabulary symbolic features, and self-construction constraints, deep learning-based methods can extract implicit and ambiguous causality. Therefore, the use of these methods for extracting causality have become widespread.

Kruengkrai et al. [6] proposed using a convolutional neural network (CNN) to extract background knowledge from noisy text to classify causality. Li et al. [7] proposed a knowledge-oriented CNN, which combines the prior knowledge in the vocabulary knowledge base to classify causality; Dasgupta et al. [8] adopted a method based on long short-term memory (LSTM). The deep model of the network determines the linguistic expression of causality in the text from the perspective of language. Misawa et al. [10] proposed the causal relationship extractor called self-attentive BiLSTM-CRF with flair embeddings (SCIFI) based on the bidirectional LSTM-BiLSTM-CRF model, using a new causal relationship labeling scheme to directly extract causal relationships.

3 Task Description and Model

3.1 Corpus

We used web scraper to collect 1000 social news items from Google Japanese news, Google English news, and Baidu news contained on WordA (ten issues of "COVID-19," "climate change," "water scarcity," "human overpopulation," "alcohol," "population aging," "urbanization," "plastic bag," "bullying," "overgrazing"). We then used crowdsourcing to artificially determine the results in the news due to WordA and summarize the answers due to these issues with appropriate terms. If the news did not directly express the cause–effect relationship regarding WordA, the answer was [none]. Excluding special answers, we collected 740 items in English, 730 in Japanese and 690 in Chinese.

3.2 Model

In this experiment, we used GPT-3, which is currently hosted online and can be accessed in real time by paying users to be allowed to query their APIs. GPT-3 is applied without any gradient updates or fine-tuning and with tasks and few shot demonstrations specified purely via text interaction with the model. GPT-3 performs well on many NLP datasets, including question answering. Therefore, we used a question–answer model to extract results generated by WordA in the new items.

GPT-3 requires the input of an introductory context to complete the task for previously unseen examples (Table 1). Introductory context refers to the textual description of the task and a demonstration of the task being completed.

Because GPT-3 does not require fine-tuning, the JSONL format is trained directly and three training language models are obtained, each for English, Japanese, and Chinese. The precisions of the three trained models are then tested. An example input for testing is shown in Table 2.

Table 1 Example for inputs of training

Prompt:
Sentence: The number of weathers, climate, and water extremes is increasing and will become more frequent and severe in many parts of the world as a result of climate change, said Mr. Talas. Question: What is the effect of climate change? Answer: weathers, climate, and water extremes
Sentence: Smith stated that water scarcity is the country's greatest challenge to economic growth and development and will only be further magnified by climate change and other socioeconomic challenges. Question: What is the effect of water scarcity? Answer: none

Table 2 Example for inputs of test

Prompt:
Sentence: Even as vaccination rates increase and COVID-19 cases go down, the impact of the health debt will continue to affect the health and well-being of Americans Question: What is the effect of COVID-19?
Answer:

Table 3 Precision of three languages models

	English	Japanese	Chinese
Precision_all	66	74.29	83.33
Precision_no_answer	54.62	84.62	93.33
Precision_effect	72.35	61.29	55.89

Because GPT-3 is speculative regarding news content, there are no prescribed criteria for validation, and we take the approach of making human judgments about the validation results. Since the reliability scores were calculated afterward, the required verification metric is precision. The precisions of the three trained models are listed in Table 3.

3.3 General Causal Relation

Singular causal relation is a causal relation triggered by individual events at a specific place and time. A general causal relation is a causal relation triggered in general [9]. Wikidata as an important knowledge source and should not be limited by singular causal relations but rather general causal relations that can provide knowledge of a causal structure.

The results extracted from GPT-3 are those from WordA in a specific news item, which is a singular causal relation. Such results are not logged into Wikidata [11], so we exhausted the next step of processing.

We prepared a large amount of news data, 3349 items in English, 1448 items in Japanese, and 1896 items in Chinese, and used GTP-3 for prediction. The results of "none" were deleted, and only those with answers were kept. The training procedure is shown in Fig. 1.

Multiple results inferred from one sentence are separated with ",". All answers were tested for similarity, and those with a similarity of 0.6 or more were grouped together. For example, among 762 Chinese news items related to alcohol, 601 inferred answers to the question "What is the result caused by alcohol? ." There were 900 answers when all the answers were stacked. We also repeatedly deleted single phrases such as "health," "blood vessels," and "effects," so they could not be used as "results." Thus, there were a total of 776 answers. All the data were then calculated using term frequency—inverse document frequency (TF-IDF) to calculate the similarity scores.

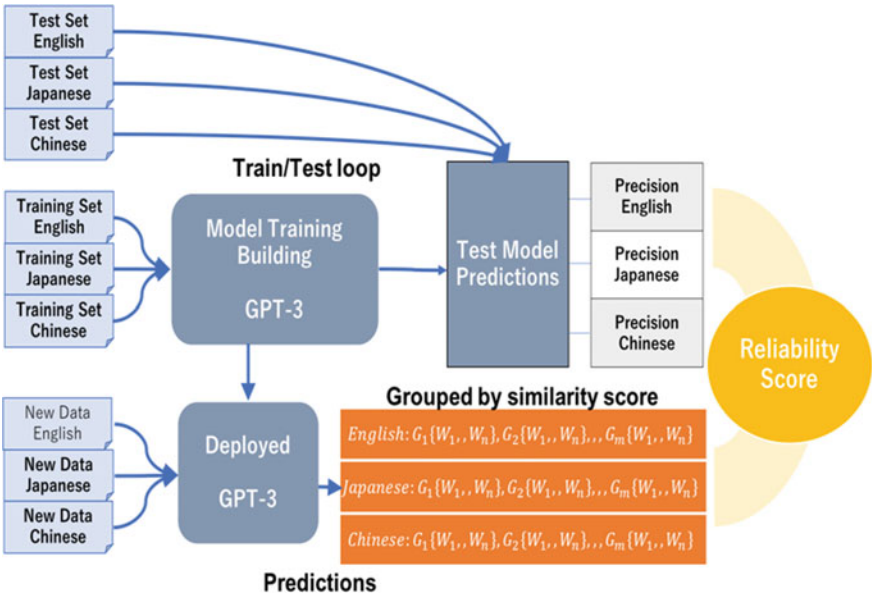


Fig. 1 Training streaming

The answers with similarity over 0.6 were grouped into one set, and 352 sets were obtained. The same method was used for all three languages to calculate the TF-IDF for each set of answers. The calculation method is as follows:

$$\mathrm{TF}_G = \frac{\text{Number of } W_i \text{ in Answers } (W_i \in \text{Group})}{\text{Number of all Answers}}$$

$$\mathrm{IDF}_G = \log \frac{\text{Number of Sentences}}{1 + \text{Number of Sentences containing the} W_i}$$

In the general results, the TF-IDF calculation is quoted, but with a slight change in the following parameters. The parameter used to calculate the TF for a group of words is the number of times all words in the group appear in the answer divided by the number of answers. The parameter used in the calculation of the IDF is the number of news sentences divided by the number of times the group of words appears in the news sentences. To increase the reliability of the predicted data, the TF-IDF of the group of words with the same meaning in the three languages are superimposed to obtain the reliability scores of the group results.

$$S_G = P_{(l|\text{total})} \sum_{l \in \{En, Ja, Ch\}} (\text{Precision}_l \times \mathrm{TF}_G \times \mathrm{IDF}_G)$$

4 Discussion

As mentioned above, we prepared 3349 items in English, 1448 items in Japanese, and 1896 items in Chinese, and used GTP-3 for prediction. Table 2 shows the high-frequency phrases in the three languages that produced outcomes due to alcohol (Table 4).

Due to language characteristics, the frequency of the results extracted from each language differed. For example, in Chinese medicine there is the statement that drinking alcohol is extraordinarily harmful to the liver, so this statement is often found on Chinese news websites, so the frequency of results due to liver damage caused by alcohol is very high in Chinese.

Using our calculations, the reliance values for each alcohol-induced outcome were calculated, yielding the top five highest values of reliance, as shown in Table 5. The results of the calculated high trust value were qualified for Wikidata by human judgment.

Table 4 TF-IDF of three models

	English	Japanese	Chinese
1	Increase risk for negative health effects, negative effects, toxics effect on body, negative effects on the body, negative effects on health 2.08	酒酔い, 二日酔い, 二日酔い, 酒酔い (drunk) 9.27	身体伤害, 伤害身体 (harmful for health) 43.70
2	Weakens immune system, negative impact on immune system 1.83	健康や日常生活に悪影響を及ぼす, 健康に悪影響を及ぼす (negative effects for health) 7.82	损害肝脏, 肝脏损伤, 危害肝脏 器官, 损害肝脏器官, 损伤肝脏, 肝脏伤害 (damage to the liver) 11.72
3	Drunk 1.66	ストレス発散 (stress reduction) 5.21	危害身体健康, 很可能会带来 不利的影响 (negative effects for health) 7.51
4	Effects on the brain, adverse effects on brain, alcohol effect on brain 1.62	女性の離婚原因の今後の増加, 離婚 (divorce) 4.12	缓解压力 (stress reduction) 4.01
5	Affect blood pressure, high blood pressure, blood pressure increases, including high blood pressure, high blood pressure 1.48	肝機能に異常をきたし (damage to the liver) 4.12	腹泻 (diarrhea) 3.7

Table 5 Reliability score of three models

	Effect of alcohol	Reliability scores
1	Negative effects on the body	14.12
2	Damage to liver	4.64
3	Drunk	3.51
4	Improves mood	2.34
5	Weakens immune system	1.85

5 Conclusion

To complement the value of the property on Wikidata, we proposed a method of automatically determining causality in text and directly extracting effect from news. The experimental results indicate that we can use this method to calculate the reliability score of the general results by extrapolating the results in multiple languages. Due to insufficient training data, prediction accuracy is not yet ideal. For future work, we will focus on semi-automatically adding values for insufficient properties on Wikidata. OpenRefine now includes a Wikidata extension that can be used to automatically add values for properties. We will also semi-automatically add estimated entities representing general causal relationships to Wikidata in bulk.

Acknowledgements This study was partially supported by JST CREST (JPMJCR15E1) and NEDO (JPNP20006).

References

1. Brown T et al (2020) Language models are few-shot learners. arXiv preprint. arXiv:2005.14165
2. Khoo CS, Kornfilt J, Oddy RN, Myaeng SH (1998) Automatic extraction of cause-effect information from newspaper text without knowledge-based inferencing. Literary Linguistic Comput 13(4):177–186
3. Khoo CS, Chan S, Niu Y (2000) Extracting causal knowledge from a medical database using graphical patterns. In: Proceedings of the 38th annual meeting of the association for computational linguistics, pp 336–343
4. Girju R, Moldovan DI (2002) Text mining for causal relations. In: FLAIRS conference, pp 360–364
5. Girju R (2003) Automatic detection of causal relations for question answering. In: Proceedings of the ACL 2003 workshop on multilingual summarization and question answering, pp 76–83
6. Kruengkrai C, Torisawa K, Hashimoto C, Kloetzer J, Oh JH, Tanaka M (2017) Improving event causality recognition with multiple background knowledge sources using multi-column convolutional neural networks. In: Proceedings of the AAAI conference on artificial intelligence, vol 31, no 1
7. Li P, Mao K (2019) Knowledge-oriented convolutional neural network for causal relation extraction from natural language texts. Expert Syst Appl 115:512–523
8. Dasgupta T, Saha R, Dey L, Naskar A (2018) Automatic extraction of causal relations from text using linguistically informed deep neural networks. In: Proceedings of the 19th annual SIGdial meeting on discourse and dialogue, pp 306–316

9. Glennan S (2011) Singular and general causal relations: a mechanist perspective. Causality Sci 789
10. Misawa S, Taniguchi M, Miura Y, Ohkuma T (2017) Character-based bidirectional LSTM-CRF with words and characters for Japanese named entity recognition. In: Proceedings of the first workshop on subword and character level models in NLP, pp 97–102
11. Allison-Cassin S, Scott D (2018) Wikidata: a platform for your library's linked open data. Code4Lib J 40

Multi-objective Evolutionary-Fuzzy for Vessel Tortuosity Characterisation

Temitope Mapayi, Pius A. Owolawi, and Adedayo O. Adio

Abstract The tortuosity characterisation of vascular networks in digital retinal fundus images plays important roles in biomedicine for the diagnosis and early detection of different human illness such as diseases of the artery and vein vessels of the retina, hypertension, and varying forms of retinopathies. Although literature findings have revealed that varying techniques have been proposed, studies have shown that there are needs for further investigation to improve the performance of automated vascular network tortuosity characterisation. This paper investigates the suitability of multi-objective evolutionary-fuzzy classification approach for the tortuosity characterisation of the vascular networks utilising the extracted geometric features of the vascular networks. The method proposed in this study seems promising as the performance accuracy rates of 88.57%, 90%, 95%, and 100% are obtained for varying training sample sizes.

Keywords Characterisation · Evolutionary-fuzzy · Networks · Retinal · Tortuosity · Vascular · Vessels

1 General Introduction

Several studies have shown that retina is a vital part of human eye utilised through various medical imaging procedures in biomedicine for the diagnosis and early detection of different human ailments such as vascular diseases and retinopathies due to high-blood pressure and diabetics [16, 21, 28]. These ailments are diagnosed by physicians using visible characteristics of the different anatomic structures such as

T. Mapayi (✉) · P. A. Owolawi
Department of Computer Systems Engineering, Faculty of Information and Communication Technology, Tshwane University of Technology, Pretoria, South Africa
e-mail: mapayit@tut.ac.za; tmapayi@yahoo.com

P. A. Owolawi
e-mail: owolawipa@tut.ac.za

A. O. Adio
Department of Ophthalmology, University of Port Harcourt, Port Harcourt, Nigeria

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_53

artery and vein vessels of the retina in the affected patients' eyes to ascertain their wellness and also keep track of their progress when undergoing treatment [16, 22]. Although the early diagnosis and efficient management of the ailments are required to prevent further worsening of the patients' health, literature findings have revealed that manual diagnostic processes required for vascular networks detection and their characterisation are laborious and time costly for the physicians, thus leading to the inefficient management of the increasing number of patients affected by these ailments [22, 25, 31]. Advancements in the development of automated methods for the detection and characterisation of the vascular networks have the capability of assisting physicians to productively and reliably manage the patients affected by the ailments [5, 19, 24, 30].

Whilst several studies have been implemented in the literature for automated vascular networks detection [1, 4, 6, 26, 29], there are indications that an highly accurate vascular network detection is one of the major fundamental requirements for an efficient vascular network tortuosity characterisation [22, 24]. Studies have also shown that vascular network tortuosity are early indications of the aforementioned ailments, even when there are no visible symptoms [7, 10, 21, 24] (see Fig. 1). The dependence of an accurate tortuosity characterisation of vascular network was considered by Bhuiyan et al. [2]. The determination of artery and vein vessels tortuosity utilising the level of convexity and concavity was implemented in [3]. Spearman correlation was applied to evaluate the set of computed vascular network tortuosity in [8, 20]. Tortuosity characterisation of arteries and veins network using a semi-automated software was implemented in [18]. AdaBoost with linear discriminant analysis was proposed in [21], and an artificial neural network technique was implemented in [22] for vessel network tortuosity characterisation. An investigation on suitability of different classifiers was presented in [23] for the tortuosity characterisation of vessel networks of the retinal. A model using chain code, Euclidean space, and curvature was implemented in [27]. The combination of chord and arc lengths with the incidence of curvatures based on stationary points was considered very effective in detecting vascular network tortuosity in [11, 12, 24]. The computation of network tortuosity using different classification models such as KNN and Naive Bayesian algorithms was implemented in [32]. Whilst diverse approaches have been

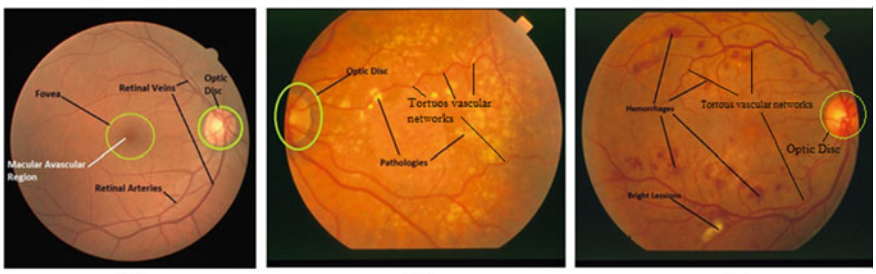


Fig. 1 **a** Anatomic structures of retinal, **b** and **c** retinal images reflecting patient's sick state

proposed in the literature for vascular network tortuosity characterisation, there are needs for additional investigation on better approaches to characterise vascular network tortuosity. This paper contributes by presenting an investigatory study on the use of multi-objective evolutionary-fuzzy (MOEF) classification approach for the tortuosity characterisation of the vessel networks. The remaining sections of this paper are: The methodology considered in this paper is described in Sect. 2. The setup of the study's experiment and the results obtained are discussed in Sect. 3, and the conclusion of this paper is drawn in Sect. 4.

2 Methodology

A detailed discussion of the multi-objective evolutionary-fuzzy (MOEF) classification approach implemented for the automatic tortuosity characterisation of the vessel networks of retina using fundus images is described in this section.

2.1 Tortuosity Characterisation of Vessel Network Using MOEF

A set of segmented vessel networks in the clinical image as show Fig. 2b are thinned, and the resulting output is preprocessed by applying a mean filter to remove noise [23]. Different geometric features comprising of normalised and non-normalised hybrid metrics with distance metric as indicated in [23] are used to characterise each of the vessel networks as either non-tortuous vessel or tortuous vessel. The usage of multi-objective evolutionary (MOE) models for fuzzy rule-based classifiers has been identified to be effective in the literature [9, 14]. This cannot, however, be farfetched from the fact that evolutionary model is effective at approximating and

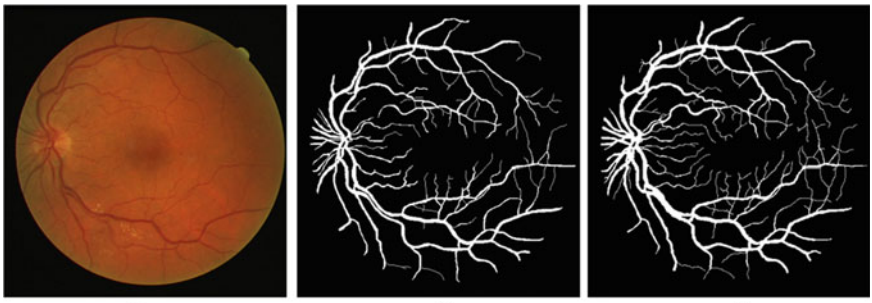


Fig. 2 **a** Sample retina image from DRIVE. **b** Manually binarized vessel networks of the image. **c** Binarized vessel network of the image serving as segmentation gold standard benchmark

optimising search problems in non-linear feature spaces that can sometimes be very complex and it will efficiently improve the output by optimising the precision of fuzzy rule-based classifiers [9, 14, 15]. Evolutionary non-dominated radial slots-based algorithm (ENORA) is an MOE algorithm based on Pareto solutions and has been identified to be efficient [9, 13, 14, 17]. This is due to the fact that an optimisation process involves the usage of certain criteria in the selection of the best features from a given set. Given that the input features from the dataset D is $x = x_1, x_2, \ldots, x_m$ and $i = 1, \ldots m$, when utilising two different objectives with no constraint, MOE is computed as

$$\begin{gathered} \text{Max./Min.} F_D(x) \\ \text{Min.} C_{\text{cal}}(x) \end{gathered} \tag{1}$$

such that $x \epsilon 1$, 0 is a boolean decision variable, where 1 indicates true when a variable is selected and 0 indicates false when a variable is not selected. $C_{\text{cal}}(x)$ is the function determines the total number of features selected, and it is computed as

$$C_{\text{cal}}(x) = \sum_{i=1}^{m} \eta(x_i) \tag{2}$$

where the function η converts true and false to 1 and 0, respectively. Given that function Υ is a fuzzy rule-based classifier and $NR_{\text{rule}}(\Upsilon)$ fuzzy rules, MOEF classifier is then computed when utilising two different objectives with four constraint as

$$\begin{aligned} \underset{\Upsilon}{\text{Max./Min.}} \quad & F_D(\Upsilon) \\ \text{Min.} \quad & \text{NR}_{\text{rule}}(\Upsilon) \\ \text{subject to:} \quad & \text{NR}_{\text{rule}}(\Upsilon) \geq n \\ & \text{NR}_{\text{rule}}(\Upsilon) \leq K_{\max} \\ & N\ell_{\text{label}}(\Upsilon) \leq L_{\min} \\ & S_{\text{sim}}(\Upsilon) \leq g_s \end{aligned} \tag{3}$$

where $F_D(\Upsilon)$ is used to measure the performance of the classifier Υ over the binarized dataset D, $n = 2$ for the number of output classes, namely non-tortuous vessel or tortuous vessel, and the function $N\ell_{label}(\Upsilon)$ described the output label. The function $S_{\text{sim}}(\Upsilon)$ also describes the fuzzy sets similarities. The maximum matching and compatibility degree as described in [14] were then implemented. The classification of the vessel networks as either non-tortuous vessel or tortuous vessel is correlated to class $C_c l$, and this is the computed as

$$f\Upsilon(x, y) = \arg_{C_c l} \max_{C_c l=1}^{n} \gamma_{C_{\text{class}}}^{\Upsilon}(x, y) \tag{4}$$

where the function $\gamma_{C_{\text{class}}}^{\Upsilon}(x, y)$ is the association degree.

3 Experimental Findings and Discussion of Results

In the experiment carried out in this study, the software used is WEKA, and the computer hardware specification is Intel Core-i5-3210M CPU with a 4 gigabyte RAM size. The dataset utilised in this paper for the experiment contains 50 digital fundus images comprising real vessel network segments each. The expert's groundtruth of the dataset was provided by an ophthalmologist consultant. The dataset is fractionated into training and test data samples. Different sizes of training data samples such as 20, 30, 40, 60, 70, and 90% are considered for training the classifier to determine if the vascular networks are tortious or not tortious. The metric used for assessing the performance of the proposed approach is the accuracy measure as indicated in Eq. 5.

$$\text{Accuracy} = (\text{TN} + \text{TP})/(\text{TN} + \text{TP} + \text{FN} + \text{FP}) \tag{5}$$

such that true positive, true negative, false positive, and false negative are denoted by TP, TN, FP, and FN, respectively.

The performance of the approach implemented in this paper and different techniques proposed in the literature is presented in Table 1, and the same dataset, as described [23], is used for all the approaches compared in the table. In comparison with AdaBoost With LDA [21] with the mean accuracy values of 67.5%, 80%, 86.67%, and 85%, the MOEF technique proposed in this paper achieved an improved classification performance of the mean accuracy values of 82.5%, 88.57%, 90%, and 95% for the training dataset sizes of 20%, 30%, 40%, and 60%, respectively, but achieved the same mean accuracy value of 100% for the training dataset sizes of

Table 1 Comparison of MOEF performance with varying approaches in the literature using different sizes of training samples

Method	Mean accuracy metric (%)						
	Training samples (%)	20 (%)	30 (%)	40 (%)	60 (%)	70 (%)	90 (%)
AdaBoost with LDA [21]		67.5	80	86.67	85	100	100
MLPNN [22]		77.5	80	83.3	85	86.67	100
KNN [23]		82.5	82.9	86.67	95	100	100
SVM [23]		37.5	71.42	73.33	75	86.67	83.3
Decision tree [23]		82.5	82.35	85.7	85	86.7	83.3
Proposed method: MOEF		82.5	88.57	90	95	100	100

70% and 90%. Although the proposed MOEF technique achieved the same mean accuracy value of 100% in comparison with MLPNN [22] for the training dataset sizes of 90%, it, however, achieved better mean accuracy values of 82.5%, 88.57%, 90%, 95%, and 100% when compared to MLPNN [22] with the mean accuracy values of 77.5%, 80%, 83.3%, 85%, and 86.67% for the training dataset sizes of 20%, 30%, 40%, 60%, and 70%, respectively. Contrasted with SVM [23] with the mean accuracy values of 37.5%, 71.42%, 73.33%, 75%, and 83.3%, a better classification performance of the mean accuracy values of 82.5%, 88.57%, 90%, 95%, 100%, and 100% is achieved by the proposed MOEF technique for the training dataset sizes of 20%, 30%, 40%, 60%, 70%, and 90%, respectively. The proposed MOEF technique achieved better classification performance of the mean accuracy values of 88.57%, 90%, 95%, 100%, and 100% is achieved in comparison with decision tree [23] with the mean accuracy values of 82.35%, 85.7%, 85%, 86.7%, and 83.3% for the training dataset sizes of 30%, 40%, 60%, 70%, and 90%, respectively. They, however, achieved the same mean accuracy value of 82.5% for the training dataset size of 20%. MOEF approach achieved higher classification performance rates with the mean accuracy values with the mean accuracy values of 88.57% and 90%, when compared to KNN [23] with the mean accuracy values of 82.9% and 86.67% for the training dataset sizes of 30% and 40%, respectively. They, however, achieved the same mean accuracy values of 82.5%, 95%, 100%, and 100% for the training dataset sizes of 20%, 60%, 70%, and 90%, respectively.

4 Conclusion

The automated tortuosity characterisation of vessel networks in digital retinal fundus images using multi-objective evolutionary-fuzzy (MOEF)classification approach on the geometric features of the vascular networks was investigated and presented in this paper. The findings of the experimental study conducted showed that the approach utilised in this study achieved promising higher mean accuracy values of 88.57%, 90%, 95%, and 100% for differing sizes of training dataset when compared with various techniques in the literature.

References

1. Aslani S, Sarnel H (2016) A new supervised retinal vessel segmentation method based on robust hybrid features. Biomed Sig Proc Cont 30:1–12
2. Bhuiyan A, Nath B, Ramamohanarao K, Kawasaki R, Wong TY (2012) Automated analysis of retinal vascular tortuosity on color retinal images. J Med Syst 36(2):689–697
3. Bribiesca E (2013) A measure of tortuosity based on chain coding. Pattern Recogn 46(3):716–724
4. Chakraborti T, Jha DK, Chowdhury AS, Jiang X (2014) A self-adaptive matched filter for retinal blood vessel detection. Mach Vis Appl 26(1):55–68

5. Chen Y (2017) A labeling-free approach to supervising deep neural networks for retinal blood vessel segmentation. arXiv preprint arXiv:1704.07502
6. Dada EG, Joseph SB (2018) Unsupervised retinal blood vessel segmentation technique using pdapso and difference image methods for detection of diabetic retinopathy. Covenant J Inf Commun Technol **6**(2)
7. Davitt BV, Wallace DK (2009) Plus disease. Survey of Ophthalmol 54(6):663–670
8. Dougherty G, Johnson MJ, Wiers MD (2010) Measurement of retinal vascular tortuosity and its application to retinal pathologies. Med Biol Eng Comput 48(1):87–95
9. Fernando J, Carlos M, Luis M, Gracia S, Sciavicco G et al (2018) Multi-objective evolutionary rule-based classification with categorical data
10. Han HC (2012) Twisted blood vessels: symptoms, etiology and biomechanical mechanisms. J Vascular Res 49(3):185–197
11. Hart WE, Goldbaum M, Côté B, Kube P, Nelson MR (1999) Measurement and classification of retinal vascular tortuosity. Int J Med Inf 53(2):239–252
12. Heneghan C, Flynn J, O'Keefe M, Cahill M (2002) Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis. Med Image Anal 6(4):407–429
13. Jimenez F, Gómez-Skarmeta AF, Sánchez G, Deb K (2002) An evolutionary algorithm for constrained multi-objective optimization. In: Proceedings of the 2002 congress on evolutionary computation. CEC'02 (Cat. No. 02TH8600). vol 2. IEEE, pp 1133–1138
14. Jiménez F, Martínez C, Marzano E, Palma JT, Sánchez G, Sciavicco G (2019) Multiobjective evolutionary feature selection for fuzzy classification. IEEE Trans Fuzzy Syst 27(5):1085–1099
15. Jiménez F, Sánchez G, Juárez JM (2014) Multi-objective evolutionary algorithms for fuzzy classification in survival prediction. Artif Intell Med 60(3):197–219
16. Joshi VS (2012) Analysis of retinal vessel networks using quantitative descriptors of vascular morphology (2012)
17. Karagoz GN, Yazici A, Dokeroglu T, Cosar A (2021) A new framework of multi-objective evolutionary algorithms for feature selection and multi-label classification of video data. Int J Mach Learn Cybernet 12(1):53–71
18. Kiely AE, Wallace DK, Freedman SF, Zhao Z (2010) Computer-assisted measurement of retinal vascular width and tortuosity in retinopathy of prematurity. Arch Ophthalmol 128(7):847–852
19. Klonoff DC, Schwartz DM (2000) An economic analysis of interventions for diabetes. Diabetes Care 23(3):390–404
20. Makkapati VV, Ravi VVC (2015) Computation of tortuosity of two dimensional vessels. In: 2015 Eighth international conference on advances in pattern recognition (ICAPR). IEEE, pp 1–4
21. Mapayi T, Dada EG, Olaifa M, Owolawi PA (2018) Retinal vessel tortuosity characterisation using boosted-linear discriminant analysis. In: 2018 International conference on intelligent and innovative computing applications (ICONIC). IEEE, pp 1–5
22. Mapayi T, Owolawi PA, Adio AO (2021) Automated detection and tortuosity characterization of retinal vascular networks. J Biomimetics, Biomater Biomed Eng 50:89–102
23. Mapayi T, Tapamo JR (2018) Performance comparison of supervised learning methods for retinal vessel tortuosity characterisation. In: 2018 conference on information communications technology and society (ICTAS), pp 1–7. IEEE
24. Mapayi T, Tapamo JR, Viriri S, Adio A (2016) Automatic retinal vessel detection and tortuosity measurement. Image Anal Stereol 35(2):117–135
25. Mapayi T, Viriri S, Tapamo JR (2015) Adaptive thresholding technique for retinal vessel segmentation based on glcm-energy information. Comput Math Methods Med 2015
26. Mapayi T, Viriri S, Tapamo JR (2015) Comparative study of retinal vessel segmentation based on global thresholding techniques. Comput Math Methods Med 2015
27. Onkaew D, Turior R, Uyyanonvara B, Akinori N, Sinthanayothin C (2011) Automatic retinal vessel tortuosity measurement using curvature of improved chain code. In: International conference on electrical, control and computer engineering 2011 (InECCE). IEEE, pp 183–186

28. Pham DL, Xu C, Prince JL (2000) Current methods in medical image segmentation 1. Ann Rev Biomed Eng 2(1):315–337
29. Roychowdhury S, Koozekanani DD, Parhi KK (2015) Iterative vessel segmentation of fundus images. IEEE Trans Biomed Eng 62(7):1738–1749
30. Soomro TA, Hellwich O, Afifi AJ, Paul M, Gao J, Zheng L (2018) Strided u-net model: retinal vessels segmentation using dice loss. In: 2018 Digital image computing: techniques and applications (DICTA). IEEE, pp 1–8
31. Staal J, Abràmoff MD, Niemeijer M, Viergever MA, van Ginneken B (2004) Ridge-based vessel segmentation in color images of the retina. Med Imag IEEE Trans 23(4):501–509
32. Turior R, Uyyanonvara B (2012) Curvature-based tortuosity evaluation for infant retinal images. J Inf Eng Appl 2(8)

AC-DC Microgrid Analysis Using a Hybrid Real-Time HiL Approach

Adrian Mihai Iuoras, Sorin Ionuț Salcu, Vasile Mihai Suciu, Lucian Nicolae Pintilie, Norbert Csaba Szekely, Mircea Bojan, and Petre Dorel Teodosescu

Abstract The paper presents a hardware-in-the-loop (HiL) simulation regarding the power flow control in an AC-DC microgrid. This microgrid topology employs the use of a combination of two bidirectional interlinking converters (BIC), with bipolar DC distribution voltage. A hybrid HiL setup is configured using the Plecs RT Box 1 and dSpace MicroLabBox 1202, on which the AC-DC microgrid hardware structure and control strategy are implemented. Observations of the microgrid's abilities to perform power factor correction/compensation (PFC) at the point of common coupling (PCC) are presented. By using the BIC, power flow control is achieved regardless of the load type and mode of operation. Also, independent control of each individual BIC is possible, this feature being highlighted in the experimental stage.

Keywords Hardware-in-the-loop (HiL) · AC-DC microgrid · Bidirectional interlinking converter (BIC) · Power flow control · Power factor correction/compensation (PFC) · Point of common coupling (PCC)

A. M. Iuoras (✉) · S. I. Salcu · V. M. Suciu · L. N. Pintilie · N. C. Szekely · M. Bojan · P. D. Teodosescu
Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, 400489 Cluj-Napoca, Romania
e-mail: adrian.iuoras@emd.utcluj.ro

S. I. Salcu
e-mail: sorin.salcu@emd.utcluj.ro

L. N. Pintilie
e-mail: lucian.pintilie@emd.utcluj.ro

N. C. Szekely
e-mail: norbert.szekely@emd.utcluj.ro

M. Bojan
e-mail: mircea.bojan@emd.utcluj.ro

P. D. Teodosescu
e-mail: petre.teodosescu@emd.utcluj.ro

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_54

1 Introduction

The increased presence of renewable energy sources brings along a set of difficulties and challenges in planning and operating the power system [1]. Stability, reliability, and power quality are mandatory aspects for the system operators, particularly due to the volatility and randomness of distributed power sources. As a way of approaching these problems effectively, the concept of using a microgrid comes into discussion. There are several types of microgrids such as DC microgrid, AC microgrid or AC-DC microgrid [2]. The last one mentioned uses a bidirectional AC-DC converter to connect the two different voltage structures of the microgrid. The study of a such converter was presented in [3]. In the mentioned paper, the focus was on the control strategy of the single-phase BIC in order to achieve load balancing and contribute to the power factor correction/compensation for an AC distribution grid, at the PCC.

Also, the concept of bipolar DC microgrid is considered. The work in [4–6] evaluates various control methods for power flow and power sharing in bipolar microgrids equipped with energy storage units, and distributed power sources. In [7, 8], the highlight has been on the interlinking converters topologies which could perform well in a bipolar DC microgrid. However, a common aspect of all these studies, is the fact that the bipolar DC voltage interlinking converter topology uses an intermediary DC-DC circuit. In the current paper, the bipolar DC voltage results from the implementation of a three-winding transformer as the input voltage source for both BIC. By connecting all BIC outputs in series, a bipolar DC voltage structure is obtained with independent control of the two voltage levels, thus with load balancing properties.

By using a HiL setup, it is possible to emulate parts of the system such as power converters or controllers using digital hardware. The aim of the setup is to reproduce different interactions of the structures as they would appear in a real system, in such a manner, that it replicates an almost precise behavior of the real physical model. As a solution for implementing the HiL setup, the Plecs RT Box 1 platform was used, along with the dSpace MicroLabBox 1202.

The most common method of implementing a HiL model using the Plecs RT Box is in conjunction with a DSP. The plant model, e.g., converter, PV system, microgrid, is being emulated on the target hardware and the control algorithm and computational process, on the TIF28069M DSP. However, in [9], a second RT Box was used for the system controller. In both situations, for the control system, these setups cannot take advantage of a direct and easy MATLAB/Simulink implementation.

In this paper, the control scheme of the HiL model was designed to run on a dSpace specific hardware solution. One of the main reasons of this approach concerning the setup is the simpler control implementation, by directly using the MATLAB/Simulink control scheme, without the need for extensive code writing and fixed-point schematic design. Furthermore, dSpace hardware is widely used in rapid-control-prototyping and control design [10, 11], while the Plecs RT Box has been optimized for hardware emulation of power converters and systems. In view of this, the main contributions

of the paper are related with the electronic power conversion structure and its steady-state validation using this new hybrid Plecs-dSpace HiL approach.

2 System Description

2.1 The Topology of the AC-DC Microgrid

The structure of the proposed system contains AC and DC circuitry delineated by two bidirectional interlinking converters, as illustrated in Fig. 1. The AC circuit contains the two BIC powered from the AC grid via a three-winding transformer with YDY connection and the LCL filters with passive damping. The DC circuits consist of two voltage levels provided by differential connection of the two AC-DC converters. The DC circuit contains a capacitive filter, the DC load, and the renewable source and/or storage element [3]. A three-phase transformer is necessary for galvanic isolation between the two BIC and in order to meet the rectifier's boost-type condition [12, 13]. The primary winding of the transformer is connected to the medium voltage AC grid and each of the two secondary windings have the voltage amplitude equal to 220 [V], being connected to the two BIC.

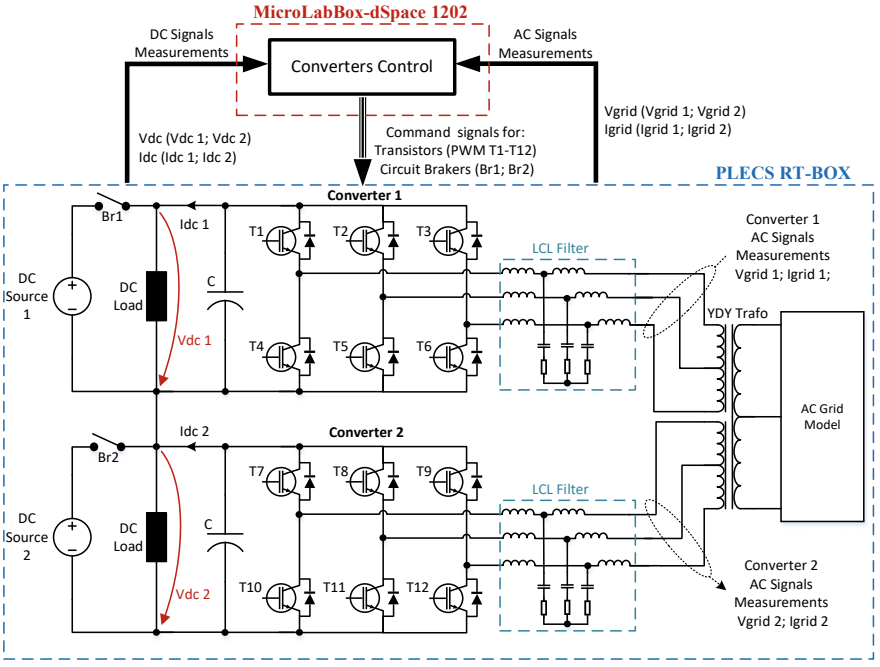


Fig. 1 Topology of the AC-DC microgrid

The bidirectional interlinking converters allow power flow in both directions according to the requirements of the energy management system. Each BIC can operate at the same time independently in rectifier or inverter mode, while contributing to the PFC of the local AC distribution public grid, at the point of common coupling [3].

The power factor correction/compensation is achieved by reactive current injection from both power converters regardless of the power flow direction. This is possible by using an appropriate control strategy based on the transfer of both active, P, and reactive, Q, power between the microgrid and the public AC grid. The AC-DC microgrid is simulated in real time through the Plecs RT Box 1 platform.

2.2 *Converters Control Strategy*

The control strategy shown in Fig. 2 is identical for both converters. It is based on the "dq synchronous reference frame" control by transforming the AC signals into DC signals for the most effective form of processing. The reference signals are the DC voltage $V_{\text{dc.ref}}$, the active power P_{ref}, which is based on Eq. (1) and the reactive power Q_{ref} requested by the AC grid management. By applying the direct three-phase dq transformation for the grid voltage, V_{grid} and grid current I_{grid}, the V_d, V_q voltages and I_d, I_q currents in dq rotation reference frame are obtained. For each converter, the measured signals for the feedback loop are the grid voltages V_{grid} and currents I_{grid} measured at the secondary transformer windings and the DC voltage V_{dc} and current I_{dc}, measured at the output of each converter [1–4].

$$P_{\text{ref}} = V_{\text{dc}} I_{\text{dc}} \tag{1}$$

The switching pattern for the two BIC is provided from two power control loops and one voltage loop. The first loop regulates the DC voltage V_{dc}, obtaining the

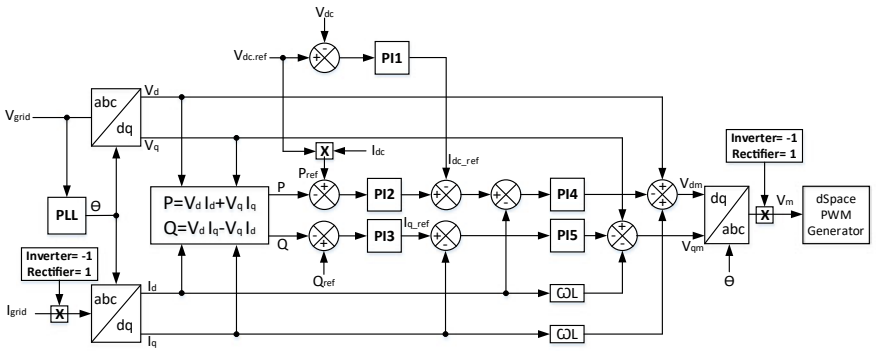


Fig. 2 Control strategy schematic

required reference for the DC current I_{DC_ref} by the proportional integral controller PI1. Through the other two loops, the active, P, and reactive, Q, power are adjusted, using as feedback the per unit outputs of P-Q computation block (Eq. 2). These signals will be compared with the references of the power control P_{ref}, Q_{ref}, and the deviation errors will be compensated by the PI2 and PI3 controllers, providing the references for two current control loops I_{d_ref}, I_{q_ref}. The signals V_{dm} and V_{qm} represent the outputs of PI4 and PI5 controllers, which after their dq to abc reference transformation will become the modulation reference voltage V_m for the dSpace PWM generator.

The rectifier or inverter operation mode is selected by shifting with 180, and° (multiplying with the constant 1) the measured grid current Igrid and the modulating signal Vm as shown in Fig. 2. The control strategies are simulated in real time through the dSpace MicroLabBox 1202 platform using the MATLAB/Simulink software [3, 14].

$$\begin{cases} P = V_d I_d + V_q I_q \\ Q = V_d I_q + V_q I_d \end{cases} \tag{2}$$

3 Implementation of HiL System Software Models

The real-time simulation models of the HiL structures were implemented using dedicated software for each hardware platforms, i.e., MATLAB/Simulink, Plecs Standalone, and control desk. These models are meant to exchange data in real time between the personal computer and the target hardware, both the modeling and control of the residential microgrid, as well as the tracking of waveforms in real time.

3.1 HiL Microgrid Modeling

The investigated microgrid model is implemented using the Plecs RT Box platform with the Plecs Standalone and Plecs Blockset dedicated software. The structure of the model is composed of two bidirectional AC-DC converters, connected to both the three-phase electrical powerline and the DC grid based on the microgrid model in Fig. 1. This model is uploaded directly to the target hardware.

3.2 Modeling the Control Structure

The control strategy modeling process for the studied residential microgrid was achieved by adapting the structure developed with the help of MATLAB/Simulink in the previous stage. In order to model the control structure, using the Micro-LabBox platform, the source code from MATLAB/Simulink is generated. This code is uploaded on the development platform via the Control Desk software.

4 Simulation Results

In this section, the real-time steady-state simulation results of the proposed microgrid model are presented. Three possible ways for manipulating the reactive energy were studied, i.e., capacitive, inductive, and resistive modes of operation, that are used in the HiL model of the microgrid.

4.1 Operation in the Rectifier Mode, with Unity Power Factor.

In this case for both converters in the structure, the reference value of the active power is imposed at 2000 [W] and the value of reactive power is 0 [VAR]. The DC voltage reference is kept constant at 350 [V]. Figure 3 shows the waveforms of the

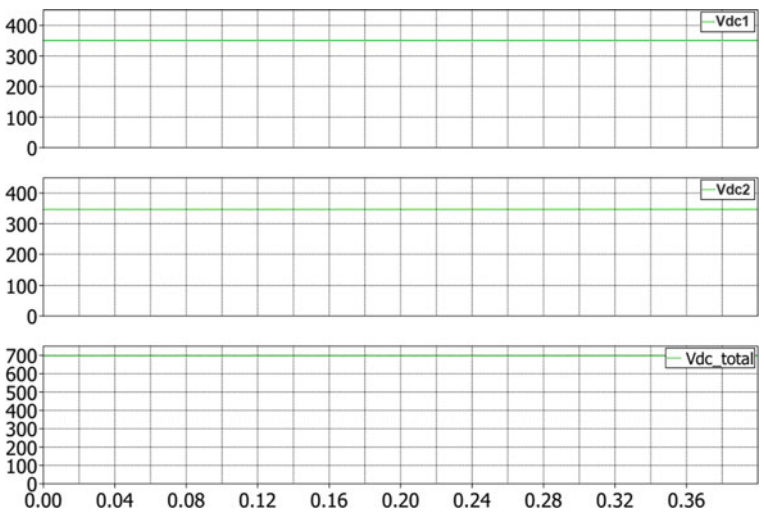


Fig. 3 Vdc1, Vdc2—voltages rectified by converters 1 and 2; Vdc_total—total rectified voltage of the two converters

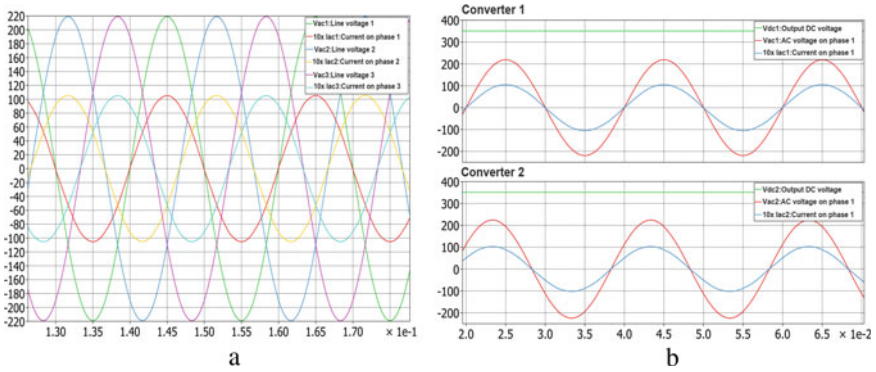


Fig. 4 Case 1. **a** three-phase currents and voltages at the filter input of the first converter; **b** current and voltage waveforms on one input phase and output rectified voltages for each converter

DC output voltage for the first converter, Vdc1, second converter, Vdc2, and the total DC output voltage, Vdc_total, (Vdc1 + Vdc2). The resulting waveforms were taken using the Plecs Standalone software.

Figure 4a shows the voltage and current waveforms at the input of the first converter's filter. In Fig. 4b, the waveforms of the voltage and current are illustrated for one phase and also the rectified voltages were depicted.

4.2 The Rectifier Mode Working Regime, with the First Converter Having a Capacitive Behavior

In this case, the reference value of active power of the first converter is imposed at 2000 [W] and the reactive power at −500 [VAR]. The second converter is working at unity power factor. The results are shown in Fig. 5a, b.

4.3 The Rectifier Mode Working Regime, with the Second Convertor Having a Capacitive Behavior

Here, the reference value of active power of the second converter is imposed at 2000 [W] and the reactive power at −500 [VAR]. First converter is working at unity power factor. The results are shown in Fig. 6a, b.

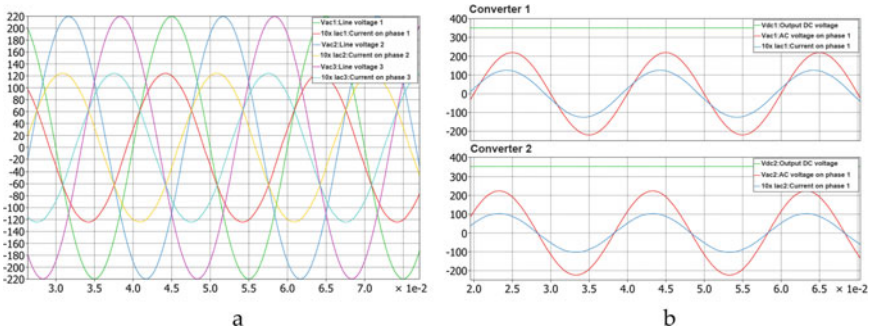


Fig. 5 Case 2. **a** three-phase currents and voltages at the filter input of the first converter; **b** current and voltage waveforms on one input phase and output rectified voltages for each converter

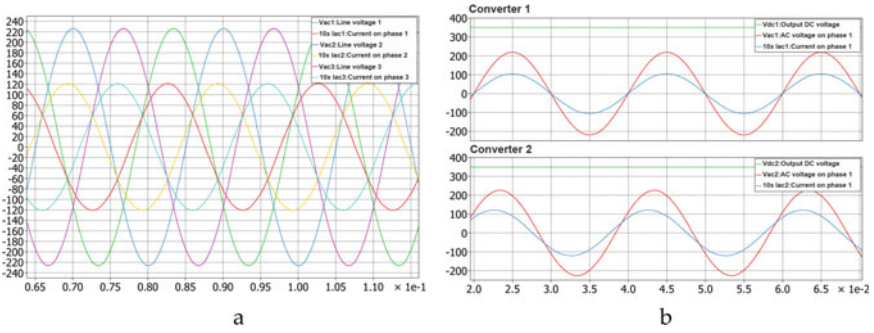


Fig. 6 Case 3. **a** three-phase currents and voltages at the filter of the second converter; **b** current and voltage waveforms on one input phase and output rectified voltages for each converter

4.4 The Rectifier Mode Working Regime, with the First Convertor Having an Inductive Behavior

In this case, the reference value of active power of the first converter is imposed at 2000 [W] and the reactive power at +500 [VAR]. Second converter is working at unity power factor. The results are shown in Fig. 7a, b.

4.5 The Rectifier Mode Working Regime, with the Second Convertor Having an Inductive Behavior

In this situation, the reference value of active power of the second converter is imposed at 2000 [W] and the reactive power at +500 [VAR]. Second converter is working at unity power factor. The results are shown in Fig. 8a, b.

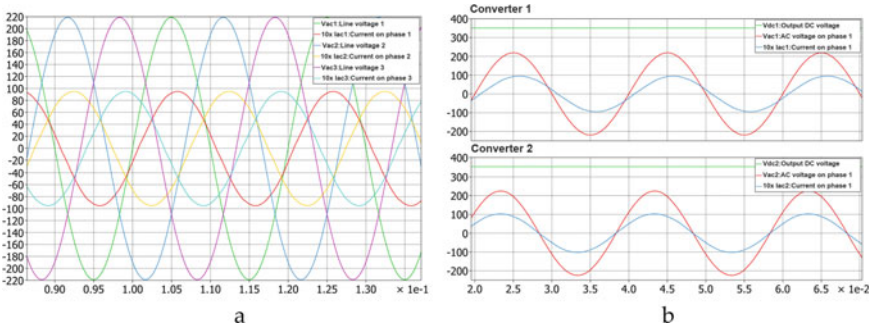


Fig. 7 Case 4. **a** three-phase currents and voltages at the input to the filter of the first converter; **b** current and voltage waveforms on one input phase and output rectified voltages for each converter

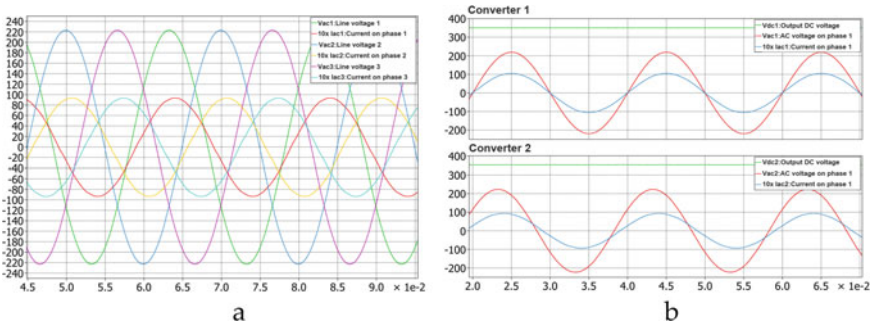


Fig. 8 Case 5—The rectifier mode working regime, with the second convertor having an inductive behavior; **a** three-phase currents and voltages at the input to the filter of the second converter; **b** current and voltage waveforms on one input phase and output rectified voltages for each converter

4.6 The First Converter Operation in the Inverter Working Regime, with Unity Power Factor

This case involves operation of the first converter in inverter mode with unity power factor and the second converter in rectifier mode. Thus, the waveforms obtained are presented in Fig. 9a, b. As it can be observed, in this mode of operation, the phase shift between voltage and current is 180°.

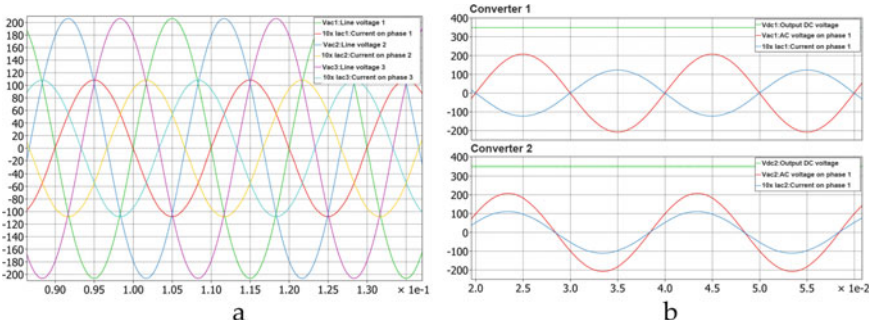


Fig. 9 Case 6—The first converter operation in the inverter working regime, with unity power factor; **a** three-phase currents and voltages at the input to the filter of the first converter; **b** current and voltage waveforms on one input phase and output rectified voltages for each converter

4.7 The First Converter Operation in the Inverter Working Regime, with Capacitive Behavior

In this situation, the first converter is operating in inverting mode having the reactive power $Q = -500$ [VAR]. The second converter is operating in rectifier mode with unity power factor. The resulting waveforms are shown in Fig. 10a, b.

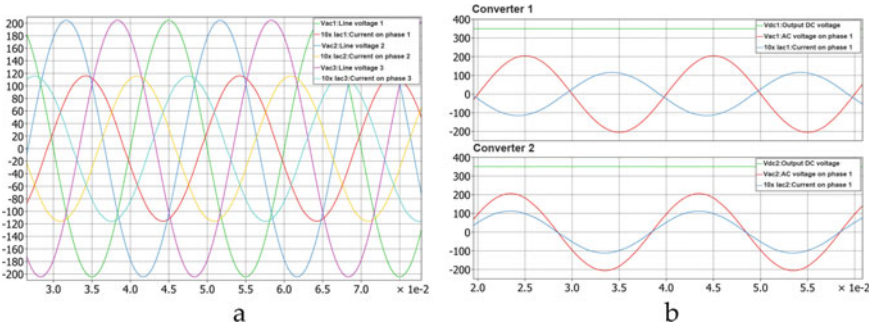


Fig. 10 Case 7—The first converter operation in the inverter working regime, with capacitive behavior; **a** three-phase currents and voltages at the filter of the second converter; **b** current and voltage waveforms on one input phase and output voltages for each converter

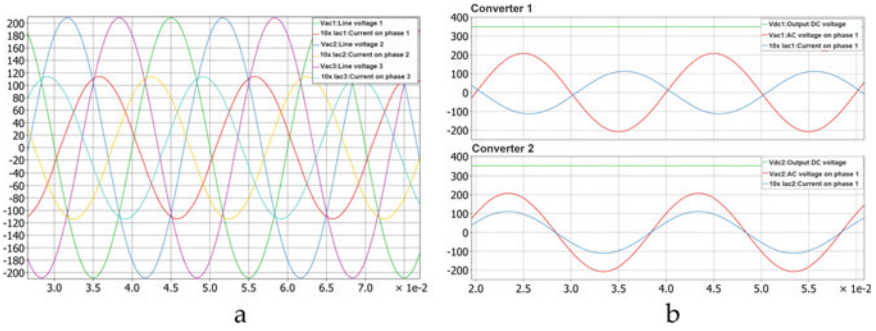


Fig. 11 Case 8—Operation in inverter mode of the first converter with inductive behavior; **a** current and voltage waveforms on one input phase and output rectified voltages for each converter; **b** current and voltage waveforms on one input phase and output voltages for each converter

4.8 Operation in Inverter Mode of the First Converter with Inductive Behavior

This case involves operation of the first converter in inverter mode having the reactive power equal to +500 [VAR] and the second converter is operating in rectifier mode with unity power factor. Thus, the waveforms obtained are presented in Fig. 11a, b.

5 Conclusions

In this paper, the HiL simulation of an AC-DC microgrid topology with two DC voltage levels was performed. The AC-DC microgrid topology is simulated in real time through the Plecs RT Box 1 platform, and the active and reactive power control strategies are implemented using the dSpace MicroLabBox 1202 platform and the MATLAB Simulink software. The novel approach of joining the two different dedicated hardware solutions is meant to take advantage of each one's better characteristics. The paper focuses on the steady-state analysis of the power management done by the electronic conversion structure and the hybrid HiL implementation. The results of the HiL simulation are presented in Sect. 4, where it is observed that the reactive power is controlled regardless of the active power flow direction, allowing the PFC to be achieved according to power requirements at the PCC of the AC distribution grid. The phase shift between voltage and current, in the presented results, is 14.47° in cases when reactive power is not equal to zero. Also, the results emphasize the capability of the proposed system to satisfy the need of a bipolar DC voltage at the microgrid level with load balancing property. Future research is expected for proper validation of the control strategy for optimal dynamic and stability behavior of the system.

Acknowledgements This paper was financially supported by the project "Entrepreneurial competences and excellence research in doctoral and postdoctoral programs—ANTREDOC," project co-funded by the European Social Fund financing agreement no. 56437/24.07.2019.

References

1. Habib A, Sou C, Hafeez HM, Arshad A (2018) Evaluation of the effect of high penetration of renewable energy sources (RES) on system frequency regulation using stochastic risk assessment technique (an approach based on improved cumulant). Renew Energy 127:204–212. https://doi.org/10.1016/j.renene.2018.04.063
2. Jiao J, Meng R, Guan Z, Ren C, Wang L, Zhang B (2019) Grid-connected control strategy for bidirectional AC-DC interlinking converter in AC-DC hybrid microgrid. In: Proceedings of the 2019 IEEE 10th international symposium on power electronics for distributed generation systems (PEDG). IEEE, Xi'an, June 2019, pp 341–345
3. Iuoras AM, Salcu SI, Rusu CG, Marginean C, Teodosescu PD (2020) Power factor compensation for a single-phase AC-DC hybrid micro-grid. In: Proceedings of the 2020 IEEE 11th international symposium on power electronics for distributed generation systems (PEDG). IEEE, Dubrovnik, Croatia, 28 Sept 2020, pp 455–458
4. Zhang C, Jiao M, Liang M, Liu M, Kan Z (2019) Power control of distributed energy storage system in bipolar DC microgrid. In: Proceedings of the 2019 22nd international conference on electrical machines and systems (ICEMS). IEEE, Harbin, China, Aug 2019, pp 1–6
5. Tavakoli SD, Mahdavyfakhr M, Hamzeh M, Sheshyekani K, Afjei E (2017) A unified control strategy for power sharing and voltage balancing in bipolar DC microgrids. Sustain Energy Grids Netw 11:58–68. https://doi.org/10.1016/j.segan.2017.07.004
6. Zhang Z, Shi D, Jin C, Koh LH, Choo FH, Wang P, Tang Y (2017) Droop control of a bipolar DC microgrid for load sharing and voltage balancing. In: Proceedings of the 2017 IEEE 3rd international future energy electronics conference and ECCE Asia (IFEEC 2017—ECCE Asia). IEEE, Kaohsiung, Taiwan, June 2017, pp 795–799
7. Park J, Byun H, Kim B, Kim S, Won C (2020) Analysis and design of coupled inductor for interleaved buck-type voltage balancer in bipolar DC microgrid. Energies 13:2775. https://doi.org/10.3390/en13112775
8. Kim J-Y, Kim H-S, Baek J-W, Jeong D-K (2019) Analysis of effective three-level neutral point clamped converter system for the bipolar LVDC distribution. Electronics 8:691. https://doi.org/10.3390/electronics8060691
9. Antoine M, Li H, Zheng X, Yu Y (2021) Multivariable state-feedback controller design for PV inverter connected to a weak grid. Energies 14:4441. https://doi.org/10.3390/en14154441
10. da Silveira Postiglione C, Simoes MG (2010) DSPACE based implementation of a grid connected smart inverter system. In: Proceedings of the 2010 IEEE 12th workshop on control and modeling for power electronics (COMPEL). IEEE, Boulder, CO, USA, June 2010, pp 1–5
11. Silva Júnior DC, Oliveira JG, de Almeida PM, Boström C (2019) Control of a multi-functional inverter in an AC microgrid—real-time simulation with control hardware in the loop. Electr Power Syst Res 172:201–212. https://doi.org/10.1016/j.epsr.2019.03.016
12. Friedli T, Hartmann M, Kolar JW (2014) The essence of three-phase PFC rectifier systems—Part II. IEEE Trans Power Electron 29:543–560. https://doi.org/10.1109/TPEL.2013.2258472
13. Kumar M, Huber L, Jovanovic MM (2014) Start-up procedure for three-phase six-switch boost PFC rectifier. In: Proceedings of the 2014 IEEE applied power electronics conference and exposition—APEC 2014. IEEE, Fort Worth, TX, USA, March 2014, pp 1852–1859
14. Tahri A, Draou A (2005) Design of a simple measuring technique of the instantaneous power in three phase system. J Electr Eng 56:221–224

Image Recognition to Detect COVID-19 Violations: Saudi Arabia Use Case

Amal Algefes, Nouf Aldossari, and Fatma Masmoudi

Abstract The upsurge in the number of criminal cases in Saudi Arabia is a cause for concern. More so, with the recent emergence of COVID-19, the government has forbidden specific social behaviors, which means that any breach of these prohibitions will be classified as a criminal. This work leverages the immense ability of deep learning architectures to develop and evaluate models to detect images of people or a person either violating or observing COVID-19 rules. For instance, an image of a person/s wearing a face mask would definitely fall under the category of non-violation, whereas an image of people hugging or shaking hands is an indication of a violation of COVID-19 rules. The model is trained and evaluated on a bunch of images that we have extracted from social media sites, and it produces exceptional results in the image classification assignment that we have performed.

Keywords Image classification · Deep learning · ResNet · Violation and non-violation · COVID-19

1 Introduction

In order to ensure a high-quality life for citizens, minimizing criminal tendencies is of high importance to the Saudi Arabia government. To do so, artificial intelligence and in particular its branches of computer vision and robotics have been absolutely pivotal in devising measures to detect, investigate, and reduce crime. Video surveillance systems have eased the work of enforcement agencies when executing security

A. Algefes (✉) · N. Aldossari · F. Masmoudi
Department of Information Systems, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Alkharj 11942, Saudi Arabia
e-mail: 441540038@std.psau.edu.sa

N. Aldossari
e-mail: 441540028@std.psau.edu.sa

F. Masmoudi
e-mail: f.masmoudi@psau.edu.sa

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_55

tasks such as facial investigation. In the context of COVID-19, The Saudi Ministry of Interior has issued legislations for the protection of citizens such as temporarily suspending Umrah for citizens and residents, suspension of international flights starting from March 15th, 2020, the suspension of studies in all schools and universities starting from March 9th, 2020. The ministry also issued fines to those not wearing a mask and no social distancing in public areas starting from May 30th, 2020. However, people still do not respect these laws, and it is too expensive in terms of effort and money to control the respect of legislation. To make the mission easier, we rely on machine learning and deep learning models on the image captured from streets, malls, parks, and so on. The importance of image processing for detecting COVID-19 violations is due to the absence of works dealing with this topic. In addition, image detection and recognition encounter several challenges due to wide variability in poses, shapes, sizes, and texture of images. In particular, facial recognition problems include posture, presence of structural components, facial emotions, and image condition. In this paper, we explore existing architectures for image recognition or classification in order to detect people committing the crime of violation of COVID-19 rules (as an extension to the previous works in the same context [1, 2]). We design a pipeline to address the problem of analyzing the images gathered pertaining to COVID-19 violations from social medial platforms. The images are pre-processed and used to train a stack of models. The rest of this paper is organized as follows. Section 2 presents the adopted methodology. Section 3 is devoted to show the results of the methodology, and Sect. 4 concludes the paper.

2 Methodology

In this section, we illustrate our methodology steps ranging from data collection to violation/non-violation images.

2.1 General Framework

We adapt the different techniques, and we use in this work to the pipeline presented in Fig. 1. First, we proceed with data collection from social media platforms. Then, we perform data preparation from images extraction to data labeling, feature extraction, and data pre-processing. Next, we train image classification through machine learning and deep learning models. Finally, we adopt models for binary classification in order to predict whether an image is violating COVID-19 rules or not.

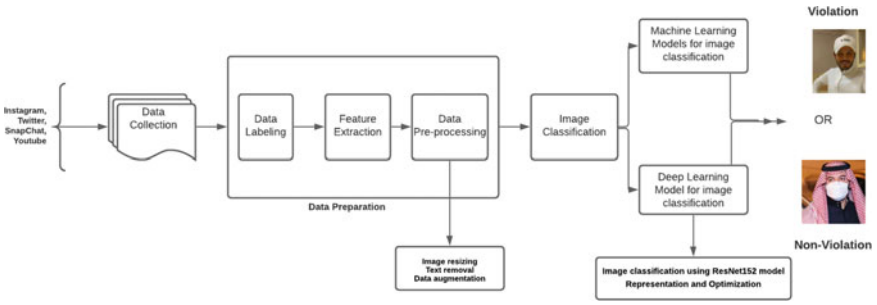


Fig. 1 Image classification pipeline to detect images of people violating COVID-19 rules

2.2 Data Collection

Prior researchers often visit the Web to query search engines to obtain images, they use in populating datasets for image classification tasks. The data collection process is responsible for many large and diverse sets of images gathering. We direct our data collection efforts to the Web in general and popular social medial applications in particular. We download images from four main applications that include Instagram, SnapChat, YouTube, and Twitter. We obtain a dataset composed of 2313 images that are taken between 2019 and 2021.

2.3 Data Preparation

In this section, we detail the data labeling, feature extraction, and data pre-processing. **Data labeling** involves assigning one (or more) class label(s) according to the concept(s) appearing in the image. For example, authors in [3] propose examples of "violin" images would be obtained by querying an image search engine for "violin" or semantically related queries such as "fiddle," "violin and piano," or "orchestra" to obtain a large and diverse set of images. Mechanical Turkers have also been used to annotate millions of images (such as the ImageNet database containing 14,197,122 images) using the semantic hierarchy of WordNet [4].

In our work, we build a binary image classification model which divides images into two classes:

- **Violation class**: It is used to represent images that depict a person or people in an act that disguises or breaks COVID-19 restrictions and rules. Examples of violation acts include not wearing a mask, shaking hands, hugging, and no social distancing in public areas.
- **Non-violation class**: It is the exact opposite of the violation class depicting people of a person observing and following COVID-19 set restrictions. Examples

Table 1 Number of images per label

Label name	Label number	Label count
Non-violation	0	666 images
Violation	1	1647 images

include wearing a face mask, social distancing in public spaces, and using elbow handshakes.

As shown in Table 1 the number of images per label as (violation and non-violation).

Feature extraction is performed to retain the most distinguishing qualities of an image. We use OpenCV[1] (a publicly available library with computer vision algorithms) to adjust the dimensions of the images and hence standardize them. It removes redundant information from the image that is irrelevant to learn, for example, nickname in SnapChat. Additionally, we filter out images that are neither jpeg nor png at this feature extraction phase.

Data pre-processing The aim of data pre-processing is to improve the quality of the collected data and make the data ready for the machine learning and deep learning model. We perform basic tasks such as resizing, color transformation, or converting *color* to *grayscale* and elimination of unclear data that are detailed in the following:

- **Image resizing**: The images are resized to a standard scale by cropping them and proportionally reducing the length and width of the images on which they are using a picture editor.
- **Text removal**: For images that containe text in them, we remove the text because it was unnecessary for our task analysis.
- **Data augmentation**: In order to increase the size of our dataset, so that we could have a lot of data to learn from, we perturb an original image by adding noise and horizontally flipping them using pythons $Scikit-learn$ package. This is performed in order to enlarge our dataset and expose the neural network to a wide variety of variations of the images. This makes it more likely that the built models recognize objects when they appear in any form and shape.

2.4 Image Classification

Various neural networks are used in deep learning for the task of image classification. In particular, convolutional neural networks (CNNs) exhibit extraordinary performance in image detection tasks because of their feature extraction ability. In this work, we adopt ResNet for image classification that alleviates the problem of vanishing gradients suffered when training very deep networks (i.e., ImageNet with

[1] https://opencv.org/.

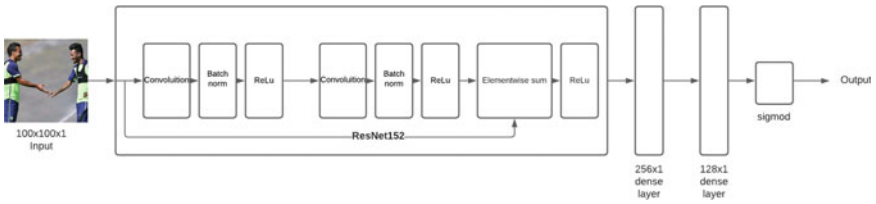


Fig. 2 ResNet architecture

16–30 CNN layers). The ResNet allows training deep neural networks with 152 layers. In the following, we explain image classification related to COVID-19 legal violations by using deep learning (CNN) and machine learning models.

1. **Deep learning for image classification**: One of the DL tasks is the extraction of disease diagnosis from clinical data, customer interests from multi-modal data in shopping applications and entertainment sites, fraudulent transactions from bank activity data, sentiments of product consumers from social media sites, and many more. In the following, we introduce image classification using the ResNet152 model and representation and optimization.

   (a) *Image classification using ResNet152 model*: Due to the computational demand of a deep layered ResNet model and similarities between images of our dataset and those in the original ImageNet dataset for which ResNet was trained, we adopt a pre-trained ResNet152 and fine-tune it for our task. The revised model shown in Fig. 2 presents pre-trained ResNet152 from TensorFlow resources to which we add three dense/full connected layers with rectified layer unit (ReLU) activation of the first two and sigmoid activation on the last.

   (b) *Representation and Optimization*: In this section, we use TensorFlow and Keras to create a sequential model. Using the Google Colab tool, we load our dataset partitioning it into 70% training and 30% testing. The network was trained based on the following parameters: RMSProp optimizer with 0.0001 learning rate and 100 epochs. We maximize resources (CPU RAM) available to the public via Google Colab.

2. **Machine learning for image classification**: In addition to the CNN-based ResNet architecture, we adopt machine learning algorithms for the image classification task using Google Colab. We adopt six algorithms including support vector machines (SVMs), K-nearest neighbors (KNNs), Naïve Bayes (NB), stochastic gradient descent classifier (SGD), decision trees (DTs), random forest (RF). The models were trained to classify the images into two labels 0 for non-violation images and 1 for images that show a violation. We load our dataset partitioning it into 70% training and 30% testing.

Table 2 Classification results when ResNet is evaluated on our dataset

Label	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC (%)
Violation (1)		98	94	96	–
Non-violation (0)		97	99	98	–
Macro-averages	97.61	98	97	97	96.6

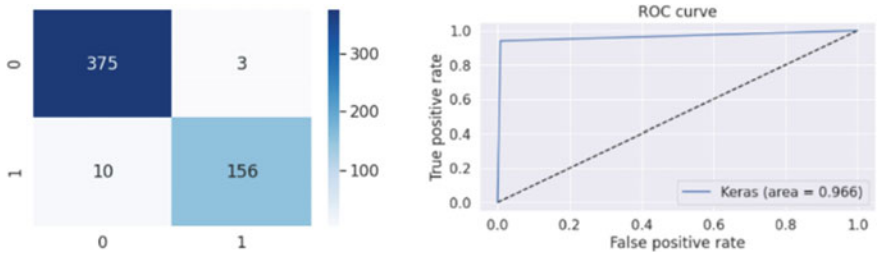


Fig. 3 ResNet result confusion matrix and ROC curve

3 Results

In this section, we present the results of both DL and ML models for images classification performed on the dataset collected in Sect. 2.

3.1 Deep Learning Results

Table 2 presents the performance evaluation results for fine-tuning ResNet152 architecture for our task. Our model performs remarkably very high scores of all five metrics computed. All metrics results are over 94%. We hypothesize that these outstanding results are due to the robustness of the ResNet model.

To obtain a better result, we adapt the confusion matrix of the ResNet model Fig. 3 As observed, TP, FP, TN, and FN are 375, 10, 156, and 3, respectively for the non-Violation label which is 0 in the figure. Similarly, TP, FP, TN, and FN are 156, 3, 375, and 10, respectively, for the violation label which takes 1 in the figure.

3.2 Machine Learning Results

Upon evaluation of the six different classification models, we obtain and report results of precision, recall, F1-score, accuracy, and AUC. Table 3 and Fig. 4 reveal RF as

Table 3 Classification results when the ML models are evaluated on our dataset

Name	Accuracy (%)	Macro precision (%)	Macro recall (%)	Macro F1-score (%)	AUC (%)
SVM	80	81	71	73	84
KNN	74	69	63	64	72
NB	69	65	66	65	69
SGD	71	68	71	69	–
DT	71	68	70	68	69
RF	82	81	75	77	85

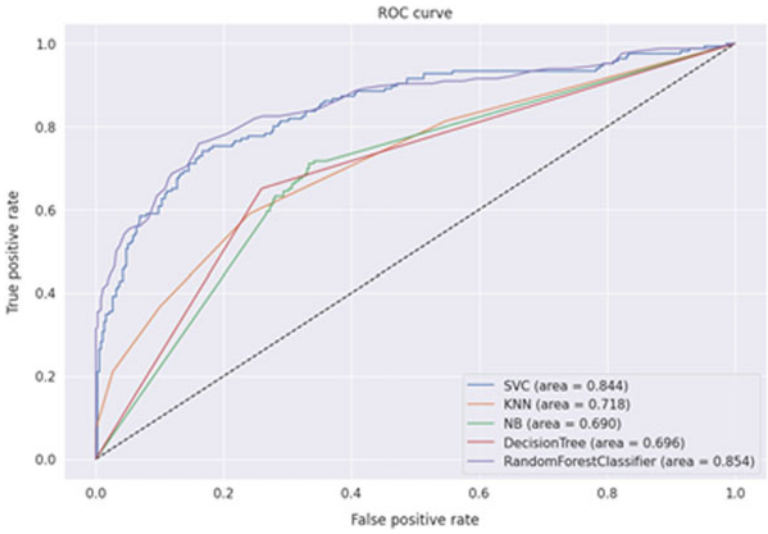


Fig. 4 ROC curves of the ML models

the best performing model outperforming all its counterparts across the five metrics computed. SVM is very competitive, too by achieving the next best performance. Despite achieving relatively high-accuracy scores, KNN, DT, SGD, and especially NB are significantly less than RF and SVM. These performance results are evidently clear in the ROC illustration in Fig. 5 RF has the largest area under its curve, and NB has the least area under its curve.

For purposes of visualizing the performance, Fig. 5 presents evaluation results of our proposed ML algorithms using accuracy metrics.

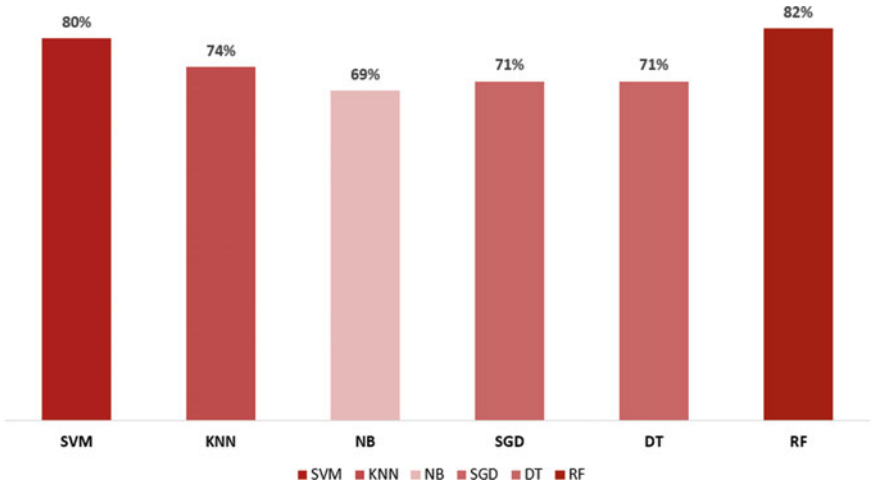


Fig. 5 Visualizing the accuracy of the ML models used for our image classification task

Table 4 Comparison between ResNet152 and best performing ML model (RF)

Model	Accuracy (%)	Macro precision (%)	Macro recall (%)	Macro F1-score (%)	AUC (%)
ResNet152	97.61	98	97	97	96.6
RF	82	81	75	77	85

3.3 Deep Learning (DL) versus Machine Learning (ML)

We compare the best performing ML model to the DL model (ResNet152) in order to determine which of the two approaches would be used to effectively achieve the task of distinguishing between violation and non-violation images. The ResNet model absolutely outperforms the RF model in every metric category as observed in Table 4.

4 Conclusion

In this paper, we present a pipeline for achieving an image classification task that involved classifying images as either violation or non-violation. Where the former implied the person or people in the image were violating COVID-19 rules, and the latter was the exact opposite. We have fine-tuned ResNet152 network architecture with a couple more fully connected layers. We additionally evaluate traditional machine learning algorithms including SVM, KNN, NB, SGD, DT, and RF for the same task. To evaluate the models, we split our dataset into a train and test set and report the

performance set. As future work, we aim at increasing the number of multi-label images, such as facilitating the procedures for wearing a mask, spacing, and others.

References

1. Algefes A, Aldossari N, Masmoudi F (2022) A text-mining approach for crime tweets in Saudi Arabia: from analysis to prediction. In: 7th international conference on data science and machine learning applications (To appear)
2. Aldossari N, Algefes A, Masmoudi F (2021) Analyzing and predicting crimes in the Kingdom of Saudi Arabia. Computand Artificial Intelligence, November 2021 in Jeddah (To appear)
3. Everingham M, Van Gool L, Williams CKI et al (2010) The PASCAL visual object classes (VOC) challenge. Int J Comput Vis 88:303–338. https://doi.org/10.1007/s11263-009-0275-4
4. Miller GA (1995) WordNet: a lexical database for English. Commun ACM 38(11):39–41. https://doi.org/10.1145/219717.219748

A Comparison of Interpretable Machine Learning Models to Predict In-Hospital Mortality After Myocardial Infarction: Analyzing Two Years Data from a High-Volume Interventional Center

Nicolai Romanov, Iolanda Valentina Popa, Alexandru Burlacu, Crischentian Brinza, and Marin Fotache

Abstract The most common cause of death among patients with cardiovascular diseases is myocardial infarction (MI). Identifying predictors for in-hospital mortality is an essential step toward MI prevention and consequent reduction in mortality. We aimed to develop machine learning (ML) methods for predicting in-hospital mortality in MI patients and apply novel techniques for models' interpretability to detect the predictive importance of the variables. Random forest (RF) and extreme gradient boosting (XGB) are applied to a dataset of 2035 MI patients who underwent percutaneous coronary intervention. When comparing the models' AUC (RF—0.9712 vs. XGB—0.9666) and accuracy (RF—97% versus XGB—98%), both techniques achieved similar performance. However, the RF model obtained a higher sensitivity (86%) than the XGBoost classifier (80%). Hypertension, cardiogenic shock, ejection fraction were identified as some of the main contributors to the outcome. Our paper contributes to the global effort of reducing mortality in patients with myocardial infarction by proposing two interpretable ML models that accurately predict in-hospital mortality in MI patients. These results are essential steps in improving current preventive strategies. However, future studies on larger datasets enriched with both categorical and continuous variables, and models validated on external data from other centers are needed to accurately assume generalizability in clinical practice.

N. Romanov (✉) · M. Fotache
"Alexandru Ioan Cuza" University, Carol I 11, 700506 Iaşi, Romania
e-mail: nicu@falcontrading.ro

M. Fotache
e-mail: fotache@uaic.ro

N. Romanov
SC Falcon Trading SRL, Iaşi, Romania

I. V. Popa · A. Burlacu · C. Brinza
"Grigore T. Popa" University of Medicine and Pharmacy, Universităţii 16, 700115 Iaşi, Romania

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_56

Keywords Machine learning · Model interpretability · Myocardial infarction · Mortality predictors

1 Introduction

The most common cause of death among patients with cardiovascular diseases is myocardial infarction (MI) [1]. MI is classified as ST-segment elevation MI (STEMI) or non-ST-segment elevation MI (NSTEMI). A significant percentage of MI cases is caused by STEMI [2], which results from the acute thrombotic occlusion of the coronary arteries that provide blood supply to the heart.

According to the latest European Society Cardiology guideline for the management of patients with STEMI, primary percutaneous coronary intervention (PCI) is the first-line treatment strategy [3]. Compared to the pharmaceutical treatment (fibrinolysis), PCI reduces mortality in acute STEMI patients by about 22% [4–7].

Although PCI proved to improve the prognosis of STEMI patients significantly, the in-hospital mortality rates after STEMI remain high (between 4–12%) [3].

Identifying predictors for in-hospital mortality is an essential step toward MI prevention and consequent reduction in mortality [3]. Our team developed and published in 2019 [8] a logistic regression (LR) model to identify independent predictors of short-term mortality after PCI. The continuous development of modern computational strategies in the area of artificial intelligence (AI)/machine learning (ML) is promising tools able to add more robustness to predictive models, increase accuracy on external validation datasets, and offer solutions for classifying the variables' importance through the models' interpretability [9–13].

This paper aimed to provide robust comparisons between the accuracy of two AI/ML methods in predicting short-term mortality after PCI, based on the two years' data from a high volume interventional center serving for the eastern part of Romania. Also, we intended to (1) identify predictors of post-PCI in-hospital mortality; (2) apply novel techniques for models' interpretability to detect the predictive importance of the variables, and (3) validate the models' on an independent dataset in a global effort to improve the management of STEMI patients and reduce residual deaths through better preventive strategies.

2 Data, Methods, and Tools

This study was conducted with a hospital-based dataset. The dataset was initially presented by Burlacu et al. [8] and contains information about 2035 patients regarding the following variables: age, gender (female = 1, male = 0), smoking status, left ventricular ejection fraction (LVEF), diabetes, hypertension, dyslipidemia, chronic kidney disease (CKD), multivessel disease (MVD) in two or more epicardial coronary arteries, proximal left anterior descending artery (pLAD) or left main artery (LM),

primary resuscitation, previous myocardial infarction (MI), cardiogenic shock, instent thrombosis, myocardial rupture, and the outcome—death. A total of 124 patients died, resulting in a 6.1% in-hospital mortality rate.

Random forest (RF) and extreme gradient boosting (XGB) are two widely used ML techniques for classification and regression tasks built using ensembles of trees. RF, presented by Breiman [14], is an example of bagging ensemble techniques that builds a series of trees by splitting a random portion of samples and features. Stable and high accuracy and low variance are some strong points of this method [15]. XGB [16] is an example of boosting techniques that build a robust model by iterating over 'weak' models and improving predictions. While RF is better at reducing model variance, XGB is known to handle the model's bias [17].

In general, the class imbalance problem solutions can be applied at the data level, model level, or both. Both models have methods to combat the problem by changing the class weights associated with the predicted labels. Since our data are highly imbalanced, with only 6.1% positive cases, during the training process, we set the RF's parameter *class_weight* to 'balanced' and XGB's *scale_pos_weigth* to 15, which is the ratio of negative samples to positive samples.

Both RF and XGB have a set of parameters, called hyper-parameters that cannot be inferred from data and need to be optimized during the training process. The hyper-parameter tunning can lead to a 5 to 20% boost in performance [18]. The following parameters were tunned for the RF model: *n_estimators* (number of trees in the model), *min_samples_split* (minimum number of samples required to split an internal node), *min_samples_leaf* (the minimum number of samples required to be at a leaf node), *max_samples* (number or ration of samples used to train a base estimator), *max_features* (number of features evaluated when looking for the best split), *max_depth* (maximum depth of a base tree), *criterion* (the function that evaluates split quality), and *class_weight* (keep default class weights or balance the data). For the XGB models, six parameters were tunned: *subsample* (ration of training instances), *min_child_weight* (minimum sum of instance weights needed in a child), *max_depth* (maximum depth of a tree), *gamma* (minimum loss reduction required to make a further partition on a leaf node of the tree), *eta* (learning rate), and *booster* (type of learner). All combinations of parameters were trained using stratified cross-validation methods to avoid the overfitting problem.

Besides obtaining a model with good performance, it is equally important to understand the factors it relies on when predicting the outcome. The task of identifying and ranking feature importance gained much interest in recent years. RF has built-in techniques that compute feature importance based on mean decrease impurity (Gini) or permutations. In this study, the Gini method was preferred. For the XGB, the available options are 'gain', 'weight', or 'cover', out of which 'gain' was chosen.

Baniecki et al. [19] proposed a set of interpretability techniques built at the model and instance levels. We used two of the instance-level method—break down (BD) and break down interactions (BDIs)—to analyze the predictions given by the RF and XGB models. The idea behind BD is to analyze the mean prediction of the model by subsequently changing all values of the features with the values from the analyzed

instance. BDI method is proper when the effects of some features depend on other features' effects, creating groups that can be interpreted together.

3 Results

This section presents the performance obtained by RF and XGB models in terms of accuracy, ROC-AUC, sensitivity, and specificity, as well as feature importance of the models and some insights into the models' predictions using different interpretability techniques.

Out of the total possible hyper-parameter combinations for each model, 1000 were randomly chosen to assess the performance using stratified cross-validation folds. The best combination of parameters for the RF model consists of the following values: n_estimators = 800, min_samples_split = 12, min_samples_leaf = 1, max_samples = 0.9, max_features = 2, max_depth = 10, criterion = gini, and class_weight = balanced. The fact that all top combinations in terms of metrics had class_weight = balanced can be considered proof that RF models can treat imbalanced data easily. Probst et al. [20] suggest that a smaller value for the max_features parameter has the advantage of building less correlated trees, thus having a more balanced aggregation. Moreover, features with moderate effect on the final prediction have greater weights if the value for max_features is low. There are two important disadvantages for having a small max_features value: The model's efficiency can be affected by the fact that the model is not built on the most optimal features, and some of the non-important features might get increased weight. For the XGB model, the parameters that yielded the best results are the following: subsample = 0.7, min_child_weight = 1, max_depth = 7, gamma = 0, eta = 0.6, scale_pos_weight = 15, and booster = gbtree. Like the RF model, XGB also had the best combinations of parameters with scale_pos_weight = 15, which handles the class imbalance problem.

The hyper-parameter tuning process was performed on a train set, subsequently divided into folds for training and validation. The final performance review of the model consists of training the model on all training data and evaluating it using test data. The metrics obtained by the model trained using hyper-parameters defined above are presented in Table 1.

Both models performed well on the test set with a score for ROC-AUC of 0.9712 and 0.9666 for RF and XGB, respectively. Both models outperformed the logistic regression (LR) used by Burlacu et al. [8] on the same dataset in terms of ROC-AUC, which was 0.933 in the initial study. The models also scored a better specificity than

Table 1 Model's metrics

Model	ROC-AUC	Accuracy	Sensitivity	Specificity
Random forest	0.9712	0.9705	0.8620	0.9775
XGBoost	0.9666	0.9803	0.8056	0.9916

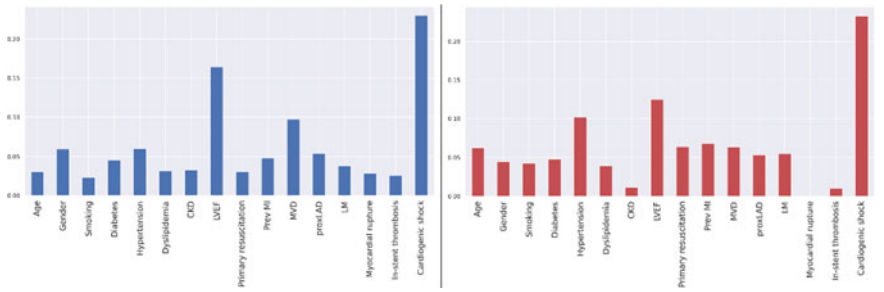


Fig. 1 Feature importance of RF (left) and XGB (right)

the LR, with 0.9775 for RF and 0.9916 for XGB compared to 0.9660 obtained initially. The sensitivity obtained in the first study was 0.8468, which is better than the one scored by XGB—0.8056 but worse than the RF—0.8620. XGB had slightly better accuracy than RF, with 0.9803 and 0.9705, respectively.

After retrieving the best parameters and training, each model's most essential features were analyzed. As shown in Fig. 1, both models ranked cardiogenic shock and LVEF as the most critical variables, followed by MVD in the RF model and hypertension in the XGB to complete the top 3. All other variables had equal importance in both models, myocardial rupture being the only one that was not considered necessary by the XGB model.

The final step of analyzing model predictions is interpreting them at the instance level. For each model, two types of plots (Break down and break down interactions) were constructed for two instances from the test set—one with true label 1, predicted label 1 (Instance 1—RF, Instance 3—XGB), and one with true label 0 and predicted label 1 (Instance 2—RF, Instance 4—XGB). These instances were chosen to identify the variables and their values that contribute toward a positive outcome of the model, whether it is correct.

As shown in Fig. 2, the main factors leading to the prediction '1' for instance 1 are the value 1 for cardiogenic shock, primary resuscitation, MVD, hypertension, proximal LAD, and gender features. Dyslipidemia and LM also added a small contribution with the value of 1.

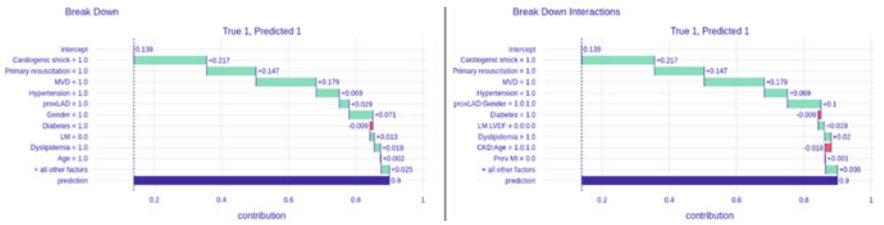


Fig. 2 Break down (left) and break down interactions (right) plots for instance 1

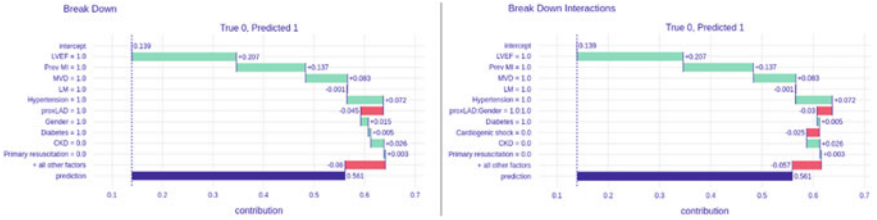


Fig. 3 Break down plot (left) and break down interactions (right) plots for instance 2

Figure 3 presents the plots for the misclassified instance. The most important features contributing to the prediction are LVEF, previous MI, LM, MVD, and hypertension; all have a value equal to 1. A small input was also added by gender = 1, diabetes = 1, and CKD = 0, while proximal LAD = 0 contributed toward lowering the probability of predicting label '1'.

The plots for the correctly classified instance by the XGB model are presented in Fig. 4. The main contributors toward the positive outcome are cardiogenic shock = 1 and LVEF = 1, with minor contributions from MVD = 1, age = 1, proximal LAD = 0, and smoking = 1.

For the mislabeled instance, presented in Fig. 5, the main contributors were cariogenic shock = 1, hypertension = 1, gender = 1, diabetes = 0, and smoking = 1. Features like LM, LVEF, and MVD with the value 0 lowered the probability of predicting a positive outcome. Overall, the two interpretability techniques are stable

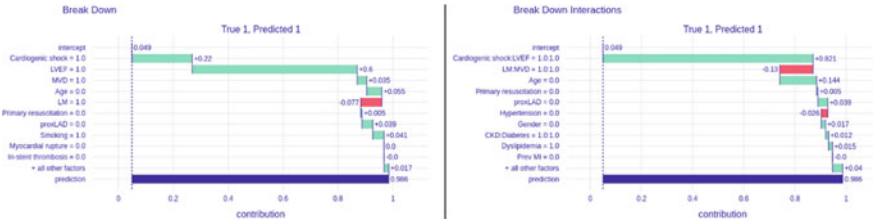


Fig. 4 Break down (left) and break down interactions (right) plots for instance 3

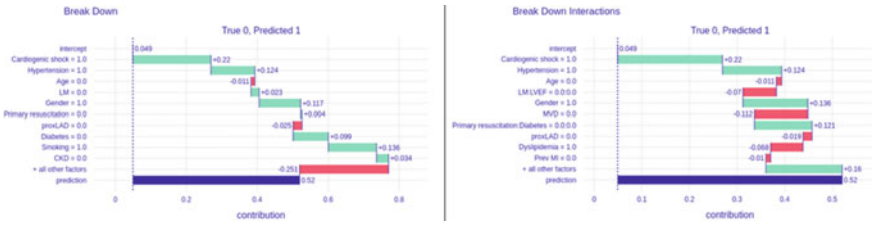


Fig. 5 Break down (left) and break down interactions (right) plots for instance 4

and consistent across instances, meaning that the same combinations of features and their values will lead to a more or less similar outcome.

4 Discussions

Our paper (1) brings novel proof for the efficacy of AI/ML methods to predict in-hospital mortality after PCI; (2) is the first research of this type using local data from our country, and (3) is in line with the global effort to reduce mortality in MI patients through the use of modern computational techniques. We provided comparisons between the accuracy of two AI/ML methods in predicting short-term mortality after PCI. We also identified predictors of post-PCI in-hospital mortality and their importance through interpretability techniques. Moreover, we validated the models' on an independent dataset.

Cardiovascular medicine has been one of the essential areas for AI/ML application. ML techniques' advantage is that they can identify patterns among vast volumes of data without the rigid assumptions of traditional statistics [21], in an era when large quantities of data are available from electronic health records and specialized medical databases.

For our study, we used two of the most popular ensemble methods for classification: random forest and boosting. When comparing the models' AUC (0.9712 vs. 0.9666) and accuracy (97 versus 98%), both techniques achieved similar performance. However, the RF model obtained a higher sensitivity (86%) than the XGBoost classifier (80%). This is relevant, as sensitivity indicates the model's ability to detect high-risk patients correctly. With high sensitivity, the algorithm will flag more patients who are actually at risk. This is important in preventive medicine to ensure the necessary preventive steps aimed to reduce mortality.

Several other AI/ML studies dedicated to predict in-hospital mortality after PCI have been conducted. Hsieh et al. [22] recorded the best performance for a decision tree (AUC = 0.895, 95% CI 0.865–0.925), outperforming other ML algorithms such as an artificial neural network (ANN), a Naïve Bayes classifier (NB), and a support vector machine (SVM). A multi-layer perceptron (MLP) model using integrated discrimination improvement and Brier score could accurately predict death after PCI (90.3% accuracy) [23]. However, the importance of the variables could not be assessed due to the lack of the model's interpretability. A random forest regression model achieved an AUC of 0.925 (95% CI 0.92 to 0.93) and found that age, congestive heart failure, and shock on presentation were leading predictors for the all-cause in-hospital mortality [11]. Sladojevic et al. [24] obtained the best prediction results for an alternating decision tree classifier that could predict in-hospital mortality with 97.8% accuracy (AUC = 0.99). They identified a subset of ten attributes most relevant to mortality prediction: urea, age, hemoglobin, DeBakey score, INR, total cholesterol, whether the patient is a physical laborer, left ventricular stroke volume, aortic velocity–time integral, and blush flow. In the paper of Al'Aref [12], a boosted ensemble algorithm (AdaBoost) had optimal discrimination with AUC of 0.927 (95%

CI 0.923–0.929) compared with AUC of 0.913 for XGBoost (95% CI 0.906–0.919, $P = 0.02$), AUC of 0.892 for RF (95% CI 0.889–0.896, $P < 0.01$), and AUC of 0.908 for logistic regression (95% CI 0.907–0.910, $P < 0.01$). Al'Aref et al. [12] identified the two most significant predictors were age and ejection fraction.

Our model has achieved better performance metrics (AUC, accuracy) than the studies published so far excepting Sladojevic et al. [24]. However, it must be noted that the studies of Kulkarni et al. [23], Zack et al. [11], and Al'Aref [12] used significantly larger datasets including more than 11,000 patients, which are essential for the demonstration of the models' generalizability.

Interestingly, any of our models did not identify age as an essential contributing factor to the outcome, unlike the studies published so far that found age to be one of the leading predictors. This may be because our dataset included predominantly elders (more than half of the patients had over 70 years of age). Moreover, hypertension was a leading contributor to the outcome by both models, but it was not mentioned in any of the other studies cited. Our paper and the referenced works identified cardiogenic shock and ejection fraction.

There are several significant limitations of this study. Firstly, the population size is smaller compared to other studies published so far, limiting the ability to assume generalizability for our models. Secondly, all the included variables are categorical, which could negatively influence the performance metrics. Thirdly, our models have been validated on an independent test set but from the same medical center. To better assess generalizability, external data are recommended for validation.

5 Conclusions

Our paper contributes to the global effort of reducing mortality in patients with myocardial infarction by proposing two interpretable AI/ML models aimed to predict in-hospital mortality after PCI and identify the outcome's predictors and their importance. These results are essential steps in improving current preventive strategies that lead to fewer preventable deaths. However, future studies on larger datasets enriched with both categorical and continuous variables and models validated on external data from other centers are needed to assume generalizability for use in clinical practice accurately.

Acknowledgements Data processing and analysis in this paper were supported by the Competitiveness Operational Program Romania, under project SMIS 124759—RaaS-IS (Research as a Service Iasi).

References

1. Collaborators GBDCOD (2018) Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 392:1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7
2. Choudhury T, West NE, El-Omar M (2016) ST elevation myocardial infarction. Clin Med (Lond) 16:277–282. https://doi.org/10.7861/clinmedicine.16-3-277
3. Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci C, Bueno H, Caforio ALP, Crea F, Goudevenos JA, Halvorsen S et al (2018) 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: the task force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). Eur Heart J 39:119–177. https://doi.org/10.1093/eurheartj/ehx393
4. Sen S, Davies JE, Malik IS, Foale RA, Mikhail GW, Hadjiloizou N, Hughes A, Mayet J, Francis DP (2012) Why does primary angioplasty not work in registries? Quantifying the susceptibility of real-world comparative effectiveness data to allocation bias. Circ Cardiovasc Qual Outcomes 5:759–766. https://doi.org/10.1161/circoutcomes.112.966853
5. Keeley EC, Boura JA, Grines CL (2003) Primary angioplasty versus intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review of 23 randomised trials. Lancet 361:13–20. https://doi.org/10.1016/s0140-6736(03)12113-7
6. Widimský P, Budesínský T, Vorác D, Groch L, Zelízko M, Aschermann M, Branny M, St'ásek J, Formánek P (2003) Long distance transport for primary angioplasty vs immediate thrombolysis in acute myocardial infarction. Final results of the randomized national multicentre trial–PRAGUE-2. Euro Heart J 24:94–104. https://doi.org/10.1016/s0195-668x(02)00468-2
7. Andersen HR, Nielsen TT, Rasmussen K, Thuesen L, Kelbaek H, Thayssen P, Abildgaard U, Pedersen F, Madsen JK, Grande P et al (2003) A comparison of coronary angioplasty with fibrinolytic therapy in acute myocardial infarction. N Engl J Med 349:733–742. https://doi.org/10.1056/NEJMoa025142
8. Burlacu A, Tinica G, Nedelciuc I, Simion P, Artene B, Covic A (2019) Strategies to lower in-hospital mortality in STEMI patients with primary PCI: analysing two years data from a high-volume interventional centre. J Interv Cardiol. https://doi.org/10.1155/2019/3402081
9. Lee W, Lee J, Woo SI, Choi SH, Bae JW, Jung S, Jeong MH, Lee WK (2021) Machine learning enhances the performance of short and long-term mortality prediction model in non-ST-segment elevation myocardial infarction. Sci Rep 11:12886. https://doi.org/10.1038/s41598-021-92362-1
10. Aziz F, Malek S, Ibrahim KS, Raja Shariff RE, Wan Ahmad WA, Ali RM, Liu KT, Selvaraj G, Kasim S (2021) Short- and long-term mortality prediction after an acute ST-elevation myocardial infarction (STEMI) in Asians: a machine learning approach. PLoS ONE 16:e0254894. https://doi.org/10.1371/journal.pone.0254894
11. Zack CJ, Senecal C, Kinar Y, Metzger Y, Bar-Sinai Y, Widmer RJ, Lennon R, Singh M, Bell MR, Lerman A et al (2019) Leveraging machine learning techniques to forecast patient prognosis after percutaneous coronary intervention. JACC Cardiovasc Interv 12:1304–1311. https://doi.org/10.1016/j.jcin.2019.02.035
12. Al'Aref SJ, Singh G, van Rosendael AR, Kolli KK, Ma X, Maliakal G, Pandey M, Lee BC, Wang J, Xu Z et al (2019) Determinants of in-hospital mortality after percutaneous coronary intervention: a machine learning approach. J Am Heart Assoc 8:e011160–e011160. https://doi.org/10.1161/JAHA.118.011160
13. Ke J, Chen Y, Wang X, Wu Z, Zhang Q, Lian Y, Chen F (2022) Machine learning-based in-hospital mortality prediction models for patients with acute coronary syndrome. Am J Emerg Med. https://doi.org/10.1016/j.ajem.2021.12.070
14. Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

15. Cutler A, Cutler DR, Stevens JR (2012) Random forests. In: Zhang C, Ma Y (eds) Ensemble machine learning: methods and applications. Springer US, Boston, MA, pp 157–175
16. Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. San Francisco, California, USA, pp 785–794
17. Efron B, Hastie T (2016) Computer age statistical inference: algorithms, evidence, and data science. Cambridge University Press, Cambridge
18. Fu W, Menzies T, Shen X (2016) Tuning for software analytics: Is it really necessary? Inf Softw Technol 76:135–146. https://doi.org/10.1016/j.infsof.2016.04.017
19. Baniecki H, Kretowicz W, Piątyszek P, Wiśniewski J, Biecek P (2021) Dalex: responsible machine learning with interactive explainability and fairness in python. J Mach Learn Res 22(214):1–7
20. Probst P, Wright MN, Boulesteix AL (2019) Hyperparameters and tuning strategies for random forest. Wiley Interdisc Rev: Data Min Knowl Disc 9:3. https://doi.org/10.1002/widm.1301
21. Breiman L (2001) Statistical modeling: the two cultures. Stat Sci 16(3):199–231
22. Hsieh MH, Lin SY, Lin CL, Hsieh MJ, Hsu WH, Ju SW, Lin CC, Hsu CY, Kao CH (2019) A fitting machine learning prediction model for short-term mortality following percutaneous catheterization intervention: a nationwide population-based study. Annal Transl Med 7:732. https://doi.org/10.21037/atm.2019.12.21
23. Kulkarni H, Amin AP (2021) Artificial intelligence in percutaneous coronary intervention: improved risk prediction of PCI-related complications using an artificial neural network. BMJ Innov 7:564–579. https://doi.org/10.1136/bmjinnov-2020-000547%JBMJInnovations
24. Sladojevic M, Pavlovic K, Velicki L, Cemerlic-Adjic N, Popov T, Tadic S, Sladojevic S, Culibrk D (2013) In-hospital mortality prediction for STEMI patients submitted to primary PCI. Euro Heart J 34.https://doi.org/10.1093/eurheartj/eht307.P478

Migration Patterns for Applications in Cloud Computing Environments

Matthias Pohl, Alexander Babel, Daniel Staegemann, Christian Haertel, Andrey Kharitonov, Abdulrahman Nahhas, and Klaus Turowski

Abstract The trend of adopting container-based systems become increasingly relevant for companies and their IT departments. Improved scalability and shorter deployment cycles in IT production are the most mentioned benefits of the technology. The adoption of container-based technologies in an existing IT system landscape requires a consideration of migration strategies. The paper at hand examines general migration patterns for the transition of virtual systems into container-based systems from a cloud computing perspectives. Several strategies are derived for a specific use case.

Keywords Cloud migration · Migration pattern · Container-based systems · Virtualised systems

M. Pohl (✉) · A. Babel · D. Staegemann · C. Haertel · A. Kharitonov · A. Nahhas · K. Turowski
Faculty of Computer Science, Otto von Guericke University, Magdeburg, Germany
e-mail: Matthias.Pohl@ovgu.de

A. Babel
e-mail: Alexander.Babel@ovgu.de

D. Staegemann
e-mail: Daniel.Staegemann@ovgu.de

C. Haertel
e-mail: Christian.Haertel@ovgu.de

A. Kharitonov
e-mail: Andrey.Kharitonov@ovgu.de

A. Nahhas
e-mail: Abdulrahman.Nahhas@ovgu.de

K. Turowski
e-mail: Klaus.Turowski@ovgu.de

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_57

1 Introduction

Container-based systems become increasingly relevant for companies and their IT departments. This significant trend is confirmed by numerous surveys and an increased search interest for technologies from this area. A study from 2020 found out that the number of companies that containerised at least 50% of their applications grew from 15 to 29% from 2018 to 2020. Meanwhile, only 21% of the companies deploy less than 10% containerised applications [1]. The Cloud Native Computing Foundation (CNCF) also conducted several surveys on the spread of container technologies in the economy. A similar picture emerges here. The use of containers for productive operations grew from 26 to 92% [9] in the period from 2016 to 2020. Another indication is the analysis of Google search queries via Google Trends. When comparing well-known container technologies such as Docker and Kubernetes with virtual machines, an increased interest in container-based solutions is apparent [15]. Consequently, container-based systems are a growth market. In 2018, revenue was approximately $1.2 billion, growing to $1.6 billion in 2019. The survey leads to an estimation of further growth to $5 billion US dollars by 2023 [26].

In addition to the trend of adopting container-based systems, the surveys provide insight into the benefits of the technology. Improved scalability and shorter deployment cycles are mentioned as the greatest impact by 51% of CNCF survey respondents [9]. Another study resumes that for 38% of respondents, high portability of the application as well as 31% ranked ease of upgrades, maintenance and lifecycle management of the application as the main advantages [17].

Whilst the advantages of migrating applications to containerised environments are well-known, the question arises as to how a system migration can be structured from a technical perspective. Since best practices are usually available for such initiatives and one would like to benefit from them, these guidelines could be the basis for a requirements analysis. The following scientific questions be formed:

RQ 1: *Which patterns can be considered for the migration of applications from a virtual to a containerised environment?*

RQ 2: *Which implementation options could be derived from the migration patterns for a specific migration use case?*

The paper at hand is structured as follows. In the next section, the process of the literature review is described and the results of the conducted review presented. Subsequently, based on the findings, suitable migration options are derived for a specific use case. A short conclusion and outlook to future work is given in the final section.

2 Migration Patterns

For further understanding, a conceptual distinction should first be made between virtual and containerised environments. A virtual machine is an environment running

an operating system including applications. It is placed on a physical system, which is called the host system or host, whilst the virtual machine is called the guest or guest system. The host system uses a hypervisor to separate the hardware from one or more isolated virtual machines. Using virtual machines on a physical system can maximise hardware utilisation [10]. A container bundles application source code, configuration files, libraries or runtime dependencies into one entity. The container can be executed on different platforms and environments without having to adapt the content. The host system only needs to provide a runtime for the container. Compared to an virtual machine, a container does not have its own operating system and instead accesses the host directly [11].

2.1 Search Results

To find patterns that can be applied for migration, it is necessary to perform a systematic literature review [20]. In the following, the review protocol is outlined and the patterns found are summarised. For the systematic literature review, the results from the search engines, Google Scholar, Web of Science (WoS) and Springer Link, are considered. The selection is based on the coverage of the databases, whilst they integrate the results of other databases such as IEEE, ACM and Elsevier, amongst others. For the literature search, at first, the term `IT system migration` is selected. Both virtualised and containerised system can be considered as IT systems. For the defined search term, Google Scholar returns around 1.2 million hits, Springer Link around 560,000 and WoS around 26,000. For manual analysis of the items, the amount is too large. Consequently, the search space is limited by specifying the term. The second search pass is performed with the term `cloud migration`. The target system will be hosted by a cloud service provider and is therefore a cloud-based system. The search term returns 535,000 entries on Google Scholar, around 38,000 on Springer Link and around 3000 on WoS. The third search uses the term "`cloud migration`". In this version, the words must occur contiguously in the respective items. The search leads to 5000 hits at Google Scholar, 444 at Springer Link and 84 at WoS. Regarding the Google Scholar result, further concretisation is necessary. In the fourth search run, the search term "`cloud migration`" `and` "`pattern`" is defined. Since patterns are to be found explicitly for a migration, the AND operation is used. The known search engines returned 1540, 208 and 5 results, respectively. Whilst the number of items from Springer Link and WoS are manually examinable, the result from Google Scholar has to be narrowed down. The fifth search uses the search term "`cloud migration`" `and` "`pattern`" `AND` ("`container`" `or` "`docker`") `AND` "`virtual machine`" and is only performed on Google Scholar, leading to 263 hits. For Springer Link, further concretisation is done by selecting two categories, Computer Science and Engineering that represent the subject area of this work. Items that fit into these categories are considered. After selecting the categories, 143 results remain for Computer Science and 32 results for Engineering.

Table 1 Search results subdivided according to search engine and search term

Search term	Springer Link	Google Scholar	WoS
IT system migration	565.855	1.230.000	26.785
Cloud migration	38.108	535.000	2.795
"Cloud migration"	444	5000	84
"Cloud migration" AND "pattern"	208	1.540	5
"Cloud migration" AND "pattern" …	55	263	–
After containment	175	133	5

The number of citations is used for the containment. A paper must have at least one citation to be analysed. This leaves 133 results. Table 1 summarises the results of the systematic search.

All contributions are then analysed by reading the abstract of the paper, as well as searching the paper itself for key terms. The keywords include `cloud`, `migration` and `pattern`. If the content of the respective work seems to fit, it is considered more closely. In addition, papers that do not address migration patterns that consider only one aspect of migration (e.g. cost) that suggest automatic migrations that focus on optimisation problems for a migration, that address at live migrations or that are strongly related to an application use case (e.g. IoT, network technologies like 5G) are furthermore excluded. Consequently, the following work summarises the result of the search (Table 2).

2.2 Analysis

A higher-level classification of patterns is proposed in a paper from 2021 [16], and integrates the cloud service models Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) as a basis [24]. It defines **Migrate to IaaS**, **Migrate to PaaS**, **Replace by SaaS**, **Revise based on SaaS** and **Reengineer to SaaS**. This classification reflects the actual migration of an application to other infrastructure or platform and the transformation of an application to a SaaS as well as a complete replacement without migration. Another widely used classification is the R-Patterns, first published by Gartner in 2010 [13]. These have been adopted by some authors [21] and by companies such as AWS [2–4, 25] or Microsoft [8]. The categories are mostly similar and have been extended over time. In the R-Pattern classification, migration to the infrastructure of another cloud service provider is called **Rehost**. In a platform migration scenario, a distinction can be made according to the effort involved. An application can be migrated to a container (**Relocate**) or with minimal platform-dependent adjustments (**Replatform** or **Revise**). More intensive

Table 2 Search results

Title	Source	Year
How to adapt applications for the Cloud environment [5]	Google Scholar	2012
How to adapt applications for the Cloud environment [5]	Springer Link	2012
CloudDSF—The cloud decision support framework for application migration [6]	Springer Link	2014
Strategies and methods for cloud migration [28]	Springer Link	2014
Cloud migration patterns: a multi-cloud service architecture perspective [18]	Springer Link	2015
Pattern-based multi-cloud architecture migration [19]	Web of science	2016
Enterprise cloud strategy [8]	Google Scholar	2016
Application migration architecture for cross clouds analysis on the strategies methods and frameworks [21]	Google Scholar	2017
Migrating to Azure: transforming legacy applications into scalable cloud-first solutions [14]	Springer Link	2018
The benefits of using experimental exploration for cloud migration analysis and planning [12]	Springer Link	2018
Microservices migration patterns [7]	Google Scholar	2018
Strategy and procedures for migration to the cloud computing [2]	Google Scholar	2018
Revisiting cloud migration: strategies and methods [27]	Google Scholar	2020
Migration—running Microsoft workloads on AWS [25]	Springer Link	2021
A multi case study on legacy system migration in the banking industry [16, 22]	Springer Link	2021
Adopting the cloud [23]	Springer Link	2021

adaptation of the application for migration to a new platform or for transformation into a SaaS is covered by the patterns **Refactor** or **Rebuild** or **Reenvision**. Replacing an application with an alternative SaaS corresponds to the pattern **Repurchase** as well as **Replace**. Finally, two cases complete the patterns. In **Retire**, an application is no longer used and therefore no migration is carried out and in **Retain**, the intention of a migration cannot generally be considered and the previous operation is maintained. A very comprehensive approach is pursued in [18]. In this framework, a distinction is made between 15 patterns that mainly relate to platform migrations and that additionally consider several detailed scenarios. The effortless transfer of an application to a new platform is marked as **MP1—Re-deployment**. The pattern **MP2—Cloudification** considers only minimal adaptations of an application through the extension of individual components with cloud services. A mixed form of MP1 and MP2 is described in the pattern **MP3—Relocation** or **MP4—Multi-Cloud Relocation**. A revision of the application, which is necessary due to the migration project, is captured in patterns for refactoring (**MP5—Multi-Cloud Refactoring**, **MP6—Hybrid refactoring**, **MP7—Hybrid refactoring with on-premise adaptation**, **MP8—Hybrid refactoring with cloud adaptation** and **MP9—Hybrid**

refactoring with hybrid adaptation). The corresponding patterns differ in the prioritisation of the target systems, i.e. whether the application is to be operated predominantly in cloud environments, on-premise or equally distributed. The patterns **MP10—Multi-Cloud Rebinding**, **MP11—Rebinding with cloud brokerage** and **MP15—Multi-Application Modernization** also address such adaptations through refactoring, but on the one hand consider the use of cloud services in communication between application components in particular and on the other hand pursue to design the application independently of the platform. Finally, the framework contains the patterns **MP12—Replacement**, **MP13—Replacement with on-premise adaptation** and **MP14—Replacement with cloud adaptation**, which intend to replace an application or some components of the application with a cloud service. Depending on whether the majority of the components are operated in the cloud environment or on-premise, a further differentiation can be made. Investigating the paper [14], the categorisation overlap with previous approaches. The transfer of an application to a target platform without adaptation is referred to as **Lift and Shift**, whilst necessary adaptations are classified as pattern **Lift, Shift, and Enhance**. The complete revision of the application due to the migration project corresponds to the pattern **Greenfield**, although there are also mixed forms, insofar as only individual parts are migrated or redeveloped and some components are even continued in the old operation (**Hybrid**). It is already evident that not all patterns of the first mentioned classifications are integrated. This also applies to the work [6] and [27]. Several previously mentioned patterns are mixed in the following patterns. The patterns **Whole migration** [27] and **Migrate the whole software stack** [6] do not precisely address whether further prerequisites are necessary for migration. A differentiated distinction between the patterns **Mixed Migration**, **Component Replacement** and **Cloud-enable Migration** [27] as well as **Partially migrate** and **Cloudify** [6] is also difficult to make. In all these patterns, only parts of the application are migrated or supplemented with cloud services. To the extent, one wants to distinguish between on-premise and cloud migration, a dividing line can be set. However, such an additional differentiation is not part of the research question. The pattern **Replace** from [6] corresponds to a complete replacement of the application with a SaaS. The other works identified from the literature search refer to the works mentioned so far and try to adapt them. Nevertheless, this does not result in any further migration approaches.

2.3 Summary

After analysing the retrieved papers and the related patterns, the summary presented in Table 3 emerges. In pattern **P1**, a complete migration of the application environment from the source system to the target system is summarised, which is carried out at the infrastructure level. A transfer to a cloud platform (PaaS) is considered in pattern **P2**. However, a further distinction is drawn as to whether no, few or enormous changes need to be made in advance so that a migration can be conducted. Additionally, the subpattern for a new development of an application for a platform is

Table 3 Summary of migration patterns

Pattern	Description	Source
P1	Hosting the application on new infrastructure without modification	Migrate to IaaS [16], Rehost [13, 21], Rehost [2–4, 25], Rehost [8], Migrate the whole software stack [6], Whole Migration [27]
P2-A	Hosting the application on a new platform without modification	Migrate to PaaS [16], Relocate [2–4, 25], Lift and Shift [14], Migrate the whole software stack [6], Whole Migration [27], MP1 [18]
P2-B	Transfer the application to a new platform with necessary modification	Migrate to PaaS [16], Replatform [2–4, 25], Revise [13, 21], Lift, Shift and Enhance [14], Migrate the whole software stack [6], Cloud-enable Migration [27], MP3 [18]
P2-C	Transfer the application to a new platform with extensive modification depending on the platform	Migrate to PaaS [16], Refactor [13, 21], Migrate the whole software stack [6], Cloud-enable Migration [27], MP3, MP4 [18]
P2-D	New development of the application to migrate to a platform	Migrate to PaaS [16], Rebuild [13, 21], Reenvision [8], Greenfield [14], Cloud-enable Migration [27], MP5, MP10, MP11, MP15 [18]
P3-A	Replacing the software by a software cloud service	Replace by SaaS [16], Replace [13, 21], Repurchase [2–4, 25], Replace [8], Replace [6], MP12 [18]
P3-B	Adaptation of the Software as a SaaS	Revised based on SaaS [16]
P3-C	New development of the application as SaaS	Reengineer to SaaS [16]
P4-A	Hybrid development, migration or replacements of components of the application to a cloud environment	Migrate to PaaS [16], Hybrid [14], MP6, MP9 [18], Mixed Migration [27]
P4-B	Extension of the current application by cloud services	Migrate to PaaS [16], Cloudify [6], MP2 [18]
P4-C	Migration of specific components of the current application to a cloud service	Migrate to PaaS [16], Partially migrate [6], Mixed Migration [27], MP7, MP8 [18]
P4-D	New development of specific components as a cloud service	Migrate to PaaS [16], Component Replacement [27], MP13, MP14 [18]
P5	The previous operation is maintained and the migration is reassessed at a later date	Retain [2–4, 25], Retain [8]
P6	Shutting down the application	Retire [2–4, 25], Retire [8]

included. The pattern **P3** mainly pursues SaaS approaches. In addition to a complete or partial redevelopment of an application as a SaaS, a replacement of the application with a SaaS can also be considered. In the pattern **P4**, approaches of hybrid application development for migration to a cloud platform are understood. The subpatterns result from the various manifestations of such hybrid migration approaches. The final patterns **P5** and **P6** consider the postponement of a migration project as well as the termination of the application.

3 Scenario Analysis

In the following, a use case is analysed that plans the migration of a PHP application, a MySQL database and an Elasticsearch cluster. The application stack is integrated in a virtual system environment in the source system. The target system should take advantage of the container-based technologies presented in Sect. 1.

The question is to be answered which implementation options result from the patterns that have been developed. First of all, a shutdown of the application (**P6**) as well as a postponement of the migration project into the future (**P5**) are not planned. A migration according to pattern **P1** would entail a complete move of the application stack to the infrastructure of a target system provided by a cloud service provider. No change to the application environment would be envisaged, so the planned container technology would not be used. The **P2** pattern provides for a migration to a PaaS. Kubernetes, for example provides a platform for container applications. However, the applications from the source system must be minimally adapted for possible migration (**P2-B**). No enormous adjustments or new developments need to be made, as containers from the technology providers (MySQL and Elasticsearch) already exist. The pattern **P3** corresponds to a simple migration to a SaaS, but this is not generally feasible due to the diversity of the application stack. However, the migration of individual components of the stack can be performed as in **P4**, which can also involve a SaaS. There are various providers of Elasticsearch clusters, so this cluster can be replaced by a SaaS (**P4-A**), but can also be migrated to the cloud by customising it (**P4-C**).

In summary, the patterns **P2-B**, as well as the patterns **P4-A** and **P4-C** can be considered.

4 Conclusion and Future Work

In the paper at hand, an overview of migration patterns for applications in cloud environments is determined with the help of a literature review. With the overview of migration patterns, individual implementation options are derived for an exemplary application scenario. This approach will be further developed in the future. A structured requirements analysis will be examined so that a detailed derivation

of the migration patterns can be achieved. Additionally, the integration of further technical parameters or performance parameters of the infrastructure and platform environments could be considered. Since the migration patterns are described largely superficially, a detailed design of the patterns is helpful so that an implementation can be carried out in a targeted manner.

References

1. 451 Research: The State of Container and Kubernetes Security. https://cyso.com/wp-content/uploads/2020/04/Stackrox-State_of_Container_and_Kubernetes_Report-winter-2020.pdf. Last accessed 27 Sep 2021
2. Ahmad N, Naveed QN, Hoda N (2018) Strategy and procedures for migration to the cloud computing. In: 2018 IEEE 5th international conference on engineering technologies and applied sciences (ICETAS)
3. Amazon web services: 6 strategies for migrating applications to the cloud. https://aws.amazon.com/blogs/enterprise-strategy/6-strategies-for-migrating-applications-to-the-cloud/. Last accessed 03 Sep 2021
4. Amazon Web Services: best practices for assessing applications to be retired during a migration to the AWS Cloud - AWS Prescriptive Guidance. https://docs.aws.amazon.com/prescriptive-guidance/latest/migration-retiring-applications/overview.html. Last accessed 05 Jan 2022
5. Andrikopoulos V, Binz T, Leymann F, Strauch S (2012) How to adapt applications for the Cloud environment. Computing 95. https://link.springer.com/article/10.1007/s00607-012-0248-2
6. Andrikopoulos V, Darsow A, Karastoyanova D, Leymann F (2014) CloudDSF–The cloud decision support framework for application migration. In: Service-oriented and cloud computing. Springer
7. Balalaie A, Heydarnoori A, Jamshidi P, Tamburri DA, Lynn T (2018) Microservices migration patterns. Softw Pract Exp https://onlinelibrary.wiley.com/doi/10.1002/spe.2608
8. Briggs B, Kassner E (2016) Enterprise cloud strategy. Pearson Educ
9. Cloud Native Computing Foundation: CNCF survey 2020. https://www.cncf.io/wp-content/uploads/2020/11/CNCF_Survey_Report_2020.pdf. Last accessed 10 Dec 2021
10. Fehling C, Leymann F, Retter R, Schupeck W, Arbitter P (2014) Cloud computing patterns. Springer Vienna, Vienna. https://doi.org/10.1007/978-3-7091-1568-8
11. Felter W, Ferreira A, Rajamony R, Rubio J (2015) An updated performance comparison of virtual machines and linux containers. In: 2015 IEEE international symposium on performance analysis of systems and software (ISPASS)
12. Fowley F, Elango DM, Magar H, Pahl C (2018) The benefits of using experimental exploration for cloud migration analysis and planning. In: Cloud computing and service science. Springer International Publishing, pp 157–176
13. Gartner: Migrating applications to the cloud: Rehost, refactor, revise, rebuild, or replace? https://www.gartner.com/en/documents/1485116/migrating-applications-to-the-cloud-rehost-refactor-revi
14. Garverick J (2018) The wireframe. In: Migrating to azure: transforming legacy applications into scalable cloud-first solutions. Apress, pp 85–103
15. Google Trends: Docker, Kubernetes, Virtual Machine. https://trends.google.com/trends/explore?date=all&q=Docker,Kubernetes,Virtual. Last accessed 10 Dec 2021
16. Hayretci HE, Aydemir FB (2021). A multi case study on legacy system migration in the banking industry. https://doi.org/10.1007/978-3-030-79382-1_32
17. IDG: 2020 cloud computing study. https://www.idg.com/tools-for-marketers/2020-cloud-computing-study/. Last accessed 10 Dec 2021

18. Jamshidi P, Pahl C, Chinenyeze S, Liu X (2015) Cloud migration patterns: a multi-cloud service architecture perspective. In: Service-oriented computing—ICSOC 2014 workshops. Springer International Publishing
19. Jamshidi P, Pahl C, Mendonça N (2016) Pattern-based multi-cloud architecture migration. Softw Pract Exp 47(9)
20. Kitchenham B, Brereton OP, Budgen D, Turner M, Bailey J, Linkman S (2009) Systematic literature reviews in software engineering–a systematic literature review. Inf Soft Technol 51(1)
21. Kumar KM, D, SK, Sardesai RP, Akhil MBSS, Kumar N (2017) Application migration architecture for cross clouds analysis on the strategies methods and frameworks. In: 2017 IEEE international conference on cloud computing in emerging markets (CCEM)
22. La Rosa M, Sadiq S, Teniente E (2021) Advanced information systems engineering–33rd international conference, CAiSE 2021. Melbourne, VIC, Australia, Proceedings. https://doi.org/10.1007/978-3-030-79382-1
23. Lisdorf A (2021) Adopting the cloud. In: Lisdorf A (ed) Cloud computing basics: a non-technical introduction. Apress, pp 169–183
24. Mell P, Grance T (2011) The NIST definition of cloud computing. NIST Special Publication 800–145
25. Pothecary R (2021) Migration—running microsoft workloads on AWS. In: Running microsoft workloads on AWS: active directory, databases, development, and more. Apress, pp 201–237. https://doi.org/10.1007/978-1-4842-6628-1_9
26. T4: Container Platform Market Share. https://www.t4.ai/industries/container-platform-market-share. Last accessed 10 Dec 2021
27. Wang Z, Yan W, Wang W (2020) Revisiting cloud migration: strategies and methods. J Phys: Conf Ser 1575. https://iopscience.iop.org/article/10.1088/1742-6596/1575/1/012232
28. Zhao JF, Zhou JT (2014) Strategies and methods for cloud migration 11(2):143–152. https://doi.org/10.1007/s11633-014-0776-7

Organization of Training in the Art Education Institution in the Context of the COVID-19 Pandemic

Maryna Tolmach, Olena Chaikovska, Svitlana Khrushch, Kateryna Kotsiubivska, and Yuliia Trach

Abstract The pandemic COVID-19 has caused significant changes in many areas of public life, including education. This article introduces how the pandemic COVID-19 affected higher education and how it can reduce the negative impacts of the pandemic and respond to future challenges with the use of ICT. Problems and consequences related to the educational process in educational institutions of Ukraine in the conditions of quarantine restrictions are analyzed. The experience of Kyiv National University of Culture and Arts in adapting the educational process to today's challenges is presented. The main directions of the university's work on reducing the consequences of the pandemic and ensuring the stability and flexibility of the educational process in the conditions of quarantine restrictions are revealed. Accented the importance of developing digital competence for all participants in the learning process.

Keywords COVID-19 · Distance learning · Digital competence

M. Tolmach (✉) · O. Chaikovska · S. Khrushch · K. Kotsiubivska · Y. Trach
Kyiv National University of Culture and Arts, Kyiv 01601, Ukraine
e-mail: margo@knukim.edu.ua

O. Chaikovska
e-mail: oachaikovska@knukim.edu.ua

S. Khrushch
e-mail: svetlanida@knukim.edu.ua

K. Kotsiubivska
e-mail: kateryna@knukim.edu.ua

Y. Trach
e-mail: trach.yuliia@knukim.edu.ua

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_58

1 Education in the Conditions of Quarantine Restrictions

1.1 Problems of Distance Learning in Higher Education Institution

Education is one of the basic human rights, which is enshrined in Article 26 of the Declaration of Human Rights, the Constitution of Ukraine and sectoral laws. The UN's fourth goal of sustainable development by 2030, which is also being implemented by Ukraine, is to "ensure comprehensive and equitable quality education and promote lifelong learning opportunities for all". The pandemic has led to significant changes in education worldwide. The vast majority of countries have closed all educational institutions for at least some time, while the rest of education has stopped in some areas. Higher education institutions have wider autonomy. Therefore, with the exception of periods of strict quarantine restrictions, they were able to independently determine the training regime depending on the infrastructure and other factors necessary to ensure the educational process in a pandemic.

The introduction of distance learning in education institutions has revealed a number of problems [4, 5].

Digital inequality. Lack of Internet access and devices needed for distance learning. In general, students from poor families are more likely to have less access to digital learning resources, such as a personal computer or laptop, high-speed Internet and a comfortable learning environment, including their own room or desk, quiet privacy and parental support. In Ukraine, access to the Internet, especially in villages and small towns, is limited, which could significantly increase educational inequalities. According to the State Statistics Committee of Ukraine, households in large cities are more likely to have access to the Internet than rural households. It should be noted the positive dynamics in the growth of the share of households that have access to Internet services at home (see Table 1).

Mental health. During a pandemic and quarantine restrictions, people may experience stress, anxiety, fear, isolation and trauma from the loss of a loved one. In Ukraine, no special studies were conducted on the mental health of educators and workers in this field, but surveys [8] were conducted on the psycho-emotional state of Ukrainians in quarantine. In the field of education, the difficulties of transition to distance learning due to lack of digital resources and skills, reduced communication and psycho-emotional support become an additional factor in the deterioration of

Table 1 Proportion of households that have access to Internet services at home

Year	In urban areas	In the countryside	All households
2019	76.3	44.1	65.8
2020	85.6	66.0	79.2

Source [1, 2]

mental health for educators, especially the older age group. In addition, the mental health of vulnerable students may also be at risk**.**

Educational outcomes. One of the negative consequences of the pandemic was the deterioration of the quality of education and educational outcomes. Much of the learning material was left for self-study. In many educational institutions, distance learning looked like sending material for self-study from a textbook, written assignments to test knowledge and evaluate results without discussion, feedback and explanations. The increase in the share of independent work and the lack of a unified approach led to an excessive workload of students. Such practices do not promote high-quality assimilation of the material, but, on the contrary, create additional stress.

Most higher education institutions remain in the mode of distance or blended learning, so the problem remains relevant and needs coordination and control by the administration of educational institutions.

Lack of systematic support for the provision of recommendations and explanations for educational institutions on educational processes during quarantine from the Ministry of Education and Science of Ukraine. During the severe quarantine period at the end of March, the Ministry of Education and Science (MESU) sent an explanation to the heads of general secondary education institutions regarding the organization of the end of the 2019/2020 academic year. The letter emphasized the autonomy of schools and decision-making on forms of distance learning at the discretion of the pedagogical council. However, the MESU did not offer guidelines and advice that teachers could follow due to lack of previous distance learning experience.

Later, the Ministry issued recommendations for the introduction of blended and distance learning in secondary and vocational education. An analysis of the experience of higher education institutions shows that despite their autonomy and the right to determine their own learning model, most institutions still needed clearer coordination at the national level to provide online learning in a pandemic.

Lack of previous distance learning experience. Most educational institutions were not ready to move to distance learning. According to a survey conducted by the State Education Quality Service in May 2020, 55.53% of schools had a problem with distance learning because they had no experience with this form of education before, so they were not ready to learn during quarantine. 47.5% of teachers (20,590 people) said that they had not previously used distance learning technology in their teaching activities [4]. A survey conducted in February 2021 showed better results but still found that most problems remain unresolved. In 2021, 47% of principals point to a lack of previous experience in implementing distance learning, and 25.5% confirm that some teachers do not have the necessary skills. 29% of pedagogical workers also noted the lack of experience in the organization of distance learning and the use of elements of distance learning in their pedagogical activities. According to a survey conducted in 2021, 34% of teachers said that they do not have enough experience in using special curricula, electronic platforms; 16% complain about the lack of experience in the effective use of computer technology [3].

1.2 Distance Learning at the Art Education

It should be emphasized that the transition to distance learning has become a challenge for the field of art education. The results of the All-Ukrainian survey "Experience of distance learning in art education" [10], conducted in April 2020 by the State Scientific and Methodological Center for the Content of Cultural and Artistic Education on the experience of distance learning and work during quarantine in art education, are noteworthy. The survey involved 2138 respondents—teachers of art schools and institutions of professional higher education, art colleges. The results demonstrate the teachers need for the development of digital skills and a fairly high level of motivation to master digital technologies.

Answers to questions about the assessment of the possibility of distance learning in art education showed that the "severity" of the perception of distance education as a possible form of learning in art school has decreased almost twice (from 41.5% before quarantining to 21.5% after quarantine). The number of its "possible supporters" also increased (from 7% to 11.5%).

The vast majority (70.5%) of art schools for the first time faced a "distance challenge", but a significant number of teachers already had experience in isolated cases (special learning conditions for students, individual subjects, educational process in colleges) apply various forms of distance learning.

A separate issue concerned the use of digital services for distance learning. In most cases, they chosen services that provide only video communication and file transfer. This suggests the need to implement a digital literacy program for teachers of art schools. Most teachers are not ready for this, neither informationally, nor methodically, nor psychologically.

Among the obstacles and difficulties in the implementation of distance learning in quarantine in art education were named various factors that can be divided into 3 categories: technical; organizational; competency [11].

Among the proposals to improve the process of implementing distance learning technologies in art education, most respondents noted the need to improve the material and technical base, the need for a single information resource to support the distance learning process and the need for digital literacy courses and in particular the use of distance learning technologies.

2 Experience of Kyiv National University of Culture and Arts During COVID-19 Pandemic

2.1 Implementing Distance Learning

Kyiv National University of Culture and Arts (KNUCA) is a modern educational corporation, national scientific and cultural center. In KNUCA, students study in such specialties: Choreography, Musical Art, Performing arts, Film and Television

art, Design, Information, Library and Archival Studies, Cultural Studies, etc. Each of the specialties has its own characteristics that must be taken into account.

KNUCA should continue teaching and learning activities, as well as involve and motivate students when there are social distancing activities. The transition to online learning can be broken down into several interrelated dimensions that affect the appropriateness and quality of distance learning provided:

a. Technical infrastructure and accessibility.
b. Digital competence of participants in the learning process and readiness to use distance learning technologies.
c. Field of study (art and classical specialties require different approaches).

2.2 Areas of Work: Challenges and Prospects

The increased use of ICT in the concept of university development. The experience gained has shown that high-quality ICT infrastructure is one of the key factors for the success of universities in ensuring the learning process in a pandemic. The use of distance learning technologies should be permanent and be one of the areas of university development. Implementation and support of the e-learning platform allow ensuring the stability and flexibility of the learning process in conditions of uncertainty.

A unified e-learning platform was created with components that ensure the following tasks:

- learning, communication and student interactions during the learning process, student identification (virtual learning environment (VLE));
- creation and improvement of educational materials (Content Management System CMS);
- learning process management (Learning management system (LMS)).

The KNUCA E-learning Platform [7] is based on the Moodle open-source platform. Moodle is a free, online Learning Management system [11] that provides: work with educational content (creation, change and storage of online courses of various formats); learning process management (personal accounts for users and teachers, training schedule); tracking the achieved learning outcomes; learning analytics tools; support for various formats for importing, exporting and migrating data; support for the Ukrainian language interface.

There are more than 6000 registered users on the platform, who have different rights to access the platform depending on the role: teachers, students, managers, deans, staff of the educational and methodological departments. A course page has been created for each discipline, which allows to place educational content, manage course participants, provide admission, verification and evaluation of tasks, and conduct various types of testing, including final testing in the discipline.

The e-course regulation has been developed for ensuring an unified approach to the development and implementation of e-learning courses. The information resource

was created for teachers with instructions and recommendations on the use of distance learning technologies, which is constantly updated. Work continues to improve the quality of e-learning materials.

A survey of students (December 2021, 1342 respondents) shows that most are satisfied with the quality of online learning (fully 56.78, partly 32.49) and provide positive feedback on disciplines where teachers actively use modern ICT. 65.8% rate the quality and content of e-materials in the Moodle system as high, 32.8%—average. Among the main advantages of online learning is the ability to work in parallel during training 69.82%, saving time 68.55%. 61.4% fully support the further application of elements of online learning, 28.5%—partially.

Providing access to video communication software for students and teachers of the university. Today, the organization of distance learning in the field of musical, choreographic and performing arts is carried out in synchronous and asynchronous modes that provide the best results, complementing each other. The main digital tool for organizing distance learning in synchronous format is video conferencing services. Important criteria for choosing a service were availability and ease of use; no time limits; number of participants up to 300 (to ensure the teaching of lectures on general subjects for faculties); high-quality sound reproduction (especially important for musical disciplines) [6, 11].

Of course, the delay in the propagation of the signal over the Internet in the synchronous format of distance learning makes it impossible for a number of educational activities that require simultaneous sound activity. Thus, regardless of the chosen service, the work of musical vocal and instrumental ensembles, choreographic ensembles, choirs and orchestras, as well as work in the conducting class is largely conducted in an abbreviated mode.

As a result of the analyzes, Google Meet and Zoom were chosen for use at the university. It should be noted that the Zoom service has an extended list of sound settings, which is critical for most teachers of creative faculties. In particular, High-Fidelity Music Mode [6] designed to provide professional-grade audio and optimize all kinds of sound and music use cases happening in Zoom Meetings. High-Fidelity Music Mode delivers professional audio from a single Zoom client, streaming to one or more listeners, for performing arts and music teachers, songwriters and anyone else looking for rich, professional-grade sound over Zoom.

KNUCA now provides training in a mixed format, combining classes face-to-face and online. To support the educational process, the university acquired Zoom for Education licenses, which allowed it to provide lectures with up to 300 participants, as well about get convenient tools for videoconferencing administration within the account and detailed usage statistics.

Infrastructure development. One of the factors of the successful provision of the educational process with the use of distance learning technologies is the availability of appropriate infrastructure. During 2020–2021, new laboratories at the Faculty of Distance Learning were opened. Computer equipment in computer laboratories has also been updated. Art studies laboratories are modernized with video communication equipment that allows communication between students and teachers at any

educational model. Increased the number of Internet access points on campus. The bandwidth of the Internet connection has been increased. In addition to the material and technical base, the staff for the Faculty of Distance Learning has been expanded. The persons responsible for the introduction of distance learning technologies have been identified at each faculty.

Planning in conditions of uncertainty. Planning the learning process in conditions of uncertainty makes new demands. In particular, a new approach to the formation of training schedules has been developed, which allows to react quickly to changes in the epidemiological situation. Additional requirements for scheduling have been introduced: planning for specific days of lectures and workshops; optimal allocation of resources (the auditorium fund, the number of licenses for video conferencing, the workload of teachers); optimization of the load on the e-learning platform during the credit-examination session.

Early planning in compliance with these requirements, makes it possible to apply different scenarios, for example, when the epidemiological situation worsens—to reduce the number of contacts by transferring lectures online, or when cases of COVID-19 are detected—providing conditions for continuing the educational process completely online for individuals academic groups or faculties.

Development of digital skills for distance and blended learning for teachers and students. The introduction of digital technologies in the educational process should be accompanied by adequate training of teachers who are ready to interact and cooperate in the latest digital environment. This has become a challenge for all parts of the art education system, for all participants in the educational process: administration, teachers, parents and students, and has highlighted the need to ensure a sufficient level of digital competence of teachers of art education institutions.

The digital skills of the participants of the educational process, first of all the scientific and pedagogical staff, were monitored, and gaps in digital skills were revealed. Department of Computer Science has carried out various measures to improve digital literacy. However, in the conditions of quarantine, taking into account the urgent need to master new digital tools, the high level of motivation of teachers and the strong support of the university administration, the trainings managed to scale and expand the audience.

Systematic measures have been organized to improve the skills of teaching staff in the field of distance and blended learning methods and to develop their digital skills. KNUCA used the international experience in the development of digital skills, in particular the experience of the Erasmus + project "dComFra—Digital competence framework for Ukrainian teachers and other citizens" (598236-EPP-1-2018-1-LT-EPPKA2-CBHE-SP). The purpose of the dComFra project is to develop a national structure of digital competencies, harmonize it with European initiatives, in particular with the European digital competence framework (DigComp), raise the digital competencies level of Ukrainian teachers and other citizens. Participation in the project activities, which took place during 2019–2020 [9], provided the best experience in organizing training from European partners, which allowed responding to

the challenges of the pandemic and effectively organizing the educational process in quarantine.

3 Conclusions

Regardless of the degree of impact of the virus on the organization, a typical situation of increased stress can occur during the following three periods: preventive measures and response to the epidemic; restoration of effective workflows; adaptation to the new normal. During each of these periods, the University took steps to address the effects of the pandemic and quarantine and to manage the risks of COVID-19 in the process.

Despite the problems associated with digital inequality, mental health, lack of skills and experience, etc., the use of ICT has provided a continuous learning process under quarantine restrictions. In particular, the experience gained helped to overcome prejudices regarding the use of digital technologies in art education; expand the range of digital technologies used in the educational process; increase the level of digital competence of all participants in the educational process; introduce new approaches using modern ICT. Further research prospects are to clarify the digital competence of culture and art teachers and find ways to improve it, formulate recommendations on the use of ICT tools for the educational process in various formats (regular, online, mixed), taking into account the characteristics of art education in a pandemic.

References

1. Access of Ukrainian households to the Internet in 2019. Statistical collection. State Statistics Service of Ukraine. https://ukrstat.org/uk/druk/publicat/kat_u/2020/zb/07/zb__dd_in19.pdf. Last accessed 30 Dec 2021
2. Access of Ukrainian households to the Internet in 2020. Statistical collection. State Statistics Service of Ukraine. http://www.ukrstat.gov.ua/druk/publicat/kat_u/2021/zb/07/zb__dd_in20.pdf. Last accessed 30 Dec 2021
3. Analytical reference on the organization of distance learning in general secondary education institutions in the 2020/2021 academic year. https://new.sqe.gov.ua/wp-content/uploads/2021/07/Analitichna-dovidka-lyutiy-2021.pdf. Last accessed 30 Dec 2021
4. Coronavirus and education: analysis of the problems and consequences of a pandemic. https://cedos.org.ua/researches/koronavirus-ta-osvita-analiz-problem-i-naslidkiv-pandemii/. Cedos, Kyiv, Ukraine (2020). Last accessed 30 Dec 2021
5. Education in a pandemic in 2020/2021: analysis of problems and consequences. https://cedos.org.ua/researches/osvita-v-umovah-pandemiyi-analiz-problem-i-naslidkiv/. Cedos, Kyiv, Ukraine (2021). Last accessed 30 Dec 2021
6. High-Fidelity, Professional-Grade Audio on Zoom, (2020). https://blog.zoom.us/high-fidelity-music-mode-professional-audio-on-zoom/. Last accessed 1 Dec 2021
7. KNUCA E-learning platform. https://elearn.knukim.edu.ua/login/index.php
8. Psycho-emotional state of Ukrainians in quarantine. http://ratinggroup.ua/research/ukraine/psihoemocionalnoe_sostoyanie_ukraincev_na_karantine.html. Last accessed 30 Dec 2021

9. Project dComFra news. http://kn.knukim.edu.ua/dcomfra-news.html. Last accessed 30 Dec 2021
10. Results of an all-Ukrainian survey of pedagogical workers in the period of quarantine: "Experience of distance learning in art education", Kyiv, Ukraine (2020). https://www.dnmczkmo.org.ua/wp-content/uploads/2020/10/anketni-danni-pdf.pdf. Last accessed 1 Dec 2021
11. Tolmach M (2021) Online learning as a challenge for art education. EDULEARN21 proceedings, pp 8868–8875. https://doi.org/10.21125/edulearn.2021.1788

An Automatic GUI Generation Method Based on Generative Adversarial Network

Xulu Yao, Moi Hoon Yap, and Yanlong Zhang

Abstract As a technique applied with artificial neural networks, deep learning is widely used in the field of image recognition. However, a lack of available datasets leads to imperfect model learning. By analysing the data scale requirements of deep learning and aiming at the application in graphical user interface (GUI) generation, it was found that the collection of GUI datasets is a time- and labour-consuming project. This makes it difficult to meet the dataset needs of current deep learning networks. To solve this problem, we propose the user interface generative adversarial network (UIGAN), a semi-supervised deep learning model, to produce a large number of reliable GUI datasets. By combining a cyclic neural network with a generated countermeasure network, UIGAN can learn the sequence relationship and characteristics of data, make the generated countermeasure network generate reasonable data, and then expand the selected Rico dataset. Relying on the network structure, the characteristics of collected data can be well analysed, and a large number of reasonable data can be generated according to them. After data processing, a reliable dataset for model training can be formed, which alleviates the problem of dataset shortage in deep learning.

Keywords GUI · Deep learning · GAN · Data augmentation

1 Introduction

The automated generation of graphical user interface (GUI) design has recently become the subject of extensive research [1, 2]. Notably, deep learning technology has become a key method for promoting advancements in this field [3, 4]. However, this method usually requires a large amount of data as support, which is a difficult problem that cannot be avoided in the implementation of automated GUI generation. Presently, it is very difficult to obtain new data directly due to related issues, such

X. Yao (✉) · M. H. Yap · Y. Zhang
Manchester Metropolitan University, Manchester M15 6BH, UK
e-mail: xulu.yao@stu.mmu.ac.uk

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_59

as object recognition in the user interface. In addition, once an image is obtained, it must be manually annotated and classified according to its layout, and these are time-consuming tasks that professionals in the field have to perform accurately [5].

Data augmentation is one of the effective methods for solving the problem of limited data at this stage [6, 7]. This technology enables the transformation of generating new training samples from an original dataset without changing the data category. The method has been successfully applied in many situations in the field of image processing, such as image classification, object recognition, semantic segmentation, and information retrieval. However, existing augmentation methods cannot be directly applied to relevant GUI design datasets due to varying data structures. In the case of natural images, the objects are usually invariant to orientation. Here, operation methods, such as rotating, tilting, and changing colours, can be used to achieve the effect of data augmentation. In the case of a GUI, the objects must be arranged in an orderly manner in accordance with the layout design specifications [8]. Therefore, there are only a limited set of operations applicable in this domain.

In this paper, we introduce a new GUI layout data augmentation method that directly synthesises a set of graphic elements in the layout. In this model, we use the Rico dataset [9] as the reference for the training data and pre-specify a set of fixed element categories (for example, "text", "button", or "image"). In our network, each element is defined by its category probability and geometric parameter representation, which are bounding box keys. The generator takes the random sampling probability of graphical elements and geometric parameters as input and arranges the chosen elements in the layout. The output includes the deterministic category and geometric parameters of the design elements, which have been chosen according to the sampling probability. The generator has the function of replacing invariants. If the input elements are re-ordered, this function will generate the same layout.

For this structured data, we implement a two-stage operation mode. The first stage directly acts on the category probability and geometric parameters of the element with size adjustments. Although effective, it is not sensitive enough for misalignments and spacing in terms of exact pixels between elements. In the second stage, we propose the user interface generative adversarial network (UIGAN) for the generator based on work in the field of vision. Just as a person can judge a design by looking at rasterised images, by mapping different elements onto a two-dimensional layout, we can evaluate their relationship with the specific methods. The models can then be used for layout optimisation because they are specifically utilised to distinguish visual patterns including, but not limited to, image segmentation and occlusion. However, the key challenge is how to map the geometric parameters to pixel level layouts. One method that can be applied is a spatial transformation network to decompose the graphic elements into bitmap masks [10].

We evaluate our method by generating a GUI layout from the frame of the markup. In summary, our model has the following contributions: (1) a differentiable wireframe generator that can determine the alignment based on the arrangement of the discrete elements and (2) UIGAN based on the generative adversarial network (GAN) [11] that directly creates structured data, which is represented in the GUI design as a set of resolution-independent markup graphic elements.

2 Related Work

2.1 Model-Based User Interface Automatic Generation

Early user interface generation methods were mainly model-based. Modellers need to fully understand the whole system and define multiple models, including task, user, presentation, session, and platform models. The models can be represented by highly professional tags to make the interface easier to create and maintain.

Puerta [11] describes a model-based interface designer (MOBI-D), a comprehensive environment that supports user-centred design through model-based interface development. To solve the problems of transitioning from scenario to formal specification and unclear UI code generation, Elkoutbi et al. [12] proposed a requirements engineering method that generates a user interface prototype from the scenario and formal specification of the application. To model interactive operation objects and realise the cooperation between interactive objects and domain objects, Silva et al. [13] designed the unified modelling language for interactive applications (UMLi), an extension of UML that provides support for UI design.

Generating graphical user interface code using machine learning technology is a relatively new research field [14, 15]. DeepCoder is a system that can generate computer programmes by using statistical prediction to enhance traditional search technology [16]. In this work, the author defines a programming language with sufficient expressiveness, including real-world programming problems. It can be predicted from input and output examples and obtains a model for mapping input and output example sets to programme attributes. Experiments were carried out showing an order of magnitude in acceleration compared with standard programme synthesis technology. This makes it possible to use this method to solve similar problems to the simplest when programming competitive websites.

In Gaunt et al. [17] research, the source code can be generated by learning the relationship between input and output examples through a differentiable interpreter. The author's goal here was to develop a new machine learning method based on a neural network and a graphical model and understand the ability of machine learning technology relative to traditional alternatives, such as the constraint-solving method based on the programming language community. The main contribution here was the proposal of TerpreT, a domain-specific language (DSL) used to express programme synthesis problems. TerpreT is similar to a probabilistic programming language: the model consists of a specification for programme representation (declaration of random variables) and an interpreter that describes how the programme maps input to output (connecting unknowns to the observed model). The reasoning task involves observing a set of input and output examples and inferring the underlying programme.

In addition, Ling et al. [18] demonstrated programme synthesis from mixed natural language and structured programme specifications as input. It is worth noting that most of these methods rely on a DSL (such as a markup, programming, or modelling language). They are designed for a specific domain, but they are usually more restrictive than a fully functional computer language. Therefore, the use of a DSL

limits the complexity of the programming language to be modelled and reduces the search space size.

There is less work to generate code through visual input-examples include hand drawing and UI screenshots. The pix2code project [19] was the first attempt to solve the problem of user interface code generated by visual input by understanding potential variables through a machine learning method rather than complex engineering heuristics. The author first generates a DSL from the prototype diagram and then compiles the DSL into source code. The author uses the design prototype map and the DSL context as training data, a convolutional neural network (CNN) to obtain image features, and two long short-term memory (LSTM) networks to understand the basic laws of the DSL contexts and the relationship between a DSL and a corresponding prototype map. On the whole, pix2code performs well, but there are some limitations, such as the need to formulate the code length range in advance and that pix2code does not consider the GUI hierarchy and code structure.

The attention-based layered decoding model of Zhu et al. [20] improved pix2code. The author proposed an attention-based code generation model, which can more finely describe GUI images and generate layered structured code consistent with the hierarchical expansion of GUI graphic elements. In addition, all the components can be extracted separately for end-to-end joint training. The experimental results show that the author's method had obvious better performance compared with the original pix2code in a public GUI code dataset and their own dataset.

Nguyen et al. [21] first proposed the technology of automatic reverse engineering of mobile application user interface (REMAUI). REMAUI automatically infers the source code of a mobile application user interface from a screenshot or conceptual design diagram of the user interface. On a given input bitmap, REMAUI identifies user interface elements through computer vision and optical character recognition (OCR) technology, infers the appropriate user interface hierarchy, and exports the results as source code for compilation and execution. The experimental evaluation results show that the UI generated by REMAUI was similar to the original UI at the pixel level and the UI hierarchy at its runtime. However, REMAUI also has limitations. First, it does not support the classification of detected components into their local component types but uses the binary classification of text or images, which limits the practical applicability of this method; second, from a developer's point of view, it is not clear whether the GUI hierarchy generated by REMAUI is really useful because the GUI hierarchy is not evaluated.

Moran et al. [22] proposed redraw based on REMAUI. In contrast to other methods, redraw is not specific to any particular field. It uses data-driven methods to classify and generate a GUI hierarchy, can use a CNN to classify GUI components into their own types, and can use a data-driven iterative k-nearest neighbours (KNN) algorithm combined with computer vision technology to generate a GUI hierarchy.

3 Methodology

3.1 Dataset

The Rico dataset can be trained with an autoencoder to understand the embedding of the GUI layout and add 64-dimensional vector annotations to each GUI to represent the encoded visual layout. The vector representation is usually used to compute structurally (and typically semantically) similar GUIs and support an example-based dataset search. To create training input for the automatic encoder embedded with layout information, we constructed a new image for each GUI, captured the bounding box area of all the leaf elements in its view hierarchy, and distinguished text, images, and buttons (Fig. 1). Rico's view hierarchy eliminates the noise image processing or OCR technology normally required to create these inputs.

3.2 Basic Layout Manipulations

The first stage of the data augmentation task includes specifying a set of transformations so that the image classification problem is considered constant, including the X-shrinking and the zoom adjustment; the transformations do not change the image category. The X-shrinking refers to the proportional shrinkage of the width of each component in each GUI layout (Fig. 2). Zoom adjustment is scaling the whole of each component in the GUI layout with the centre coordinate as the base point (Fig. 3). With these two methods, the number of Rico datasets can be doubled each time they are scaled by a certain percentage (e.g. 5%, 10%). It should be noted that the image enhancement technique depends on the problem, and certain transformations should not be applied.

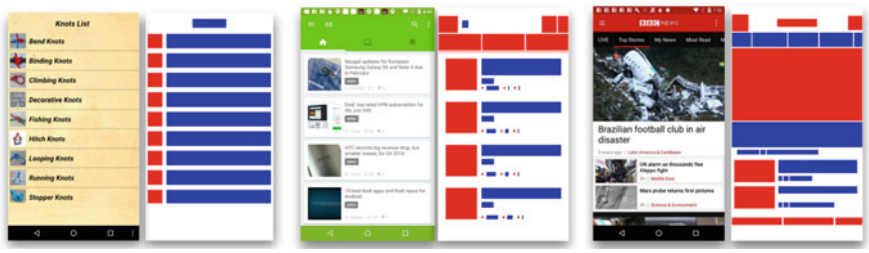


Fig. 1 GUI layouts constructed for the Rico dataset

Fig. 2 Transformation of X-shrinking

Fig. 3 Transformation of zoom adjustment

3.3 Layout Generation with GAN

The second stage is based on layout generation, in which all design elements are arranged in an appropriate size and position according to their content-based attributes (such as area, aspect ratio, and reading order). We use a generative adversarial network (GAN) to automatically generate the layouts of the Rico dataset.

A GAN [23] can create data similar to original data through the complex operation of a neural network. An excellent GAN can closely imitate the characteristics of the original data. The accuracy of its output directly affects the results of subsequent research on behaviour recognition, an outcome that is very significant to follow-up work.

GANs are widely used in various fields [24]. However, they remain very rare in the field of GUI dataset generation. By analysing the characteristics of the layout structures in the Rico dataset, we find no great difference between the layout structure data and sequence data. Thus, GANs are also suitable for the expansion of the GUI datasets. For a generating process, a special network structure is necessary to generate satisfactory layout data.

After the selection of the generator and discriminator, we propose UIGAN, a modified GAN framework especially for GUI layout generation. The structure of UIGAN proposed in this paper is shown in Fig. 4.

As seen in the architecture of UIGAN, in addition to using a non-traditional recurrent neural network (RNN) [25] and a convolutional neural network (CNN) [26], the discriminator and generator must also be connected to the fully connected layer (FC). The generator mainly consists of three layers of RNN, which are used to generate characteristic data with time correlations. Each layer of the RNN includes

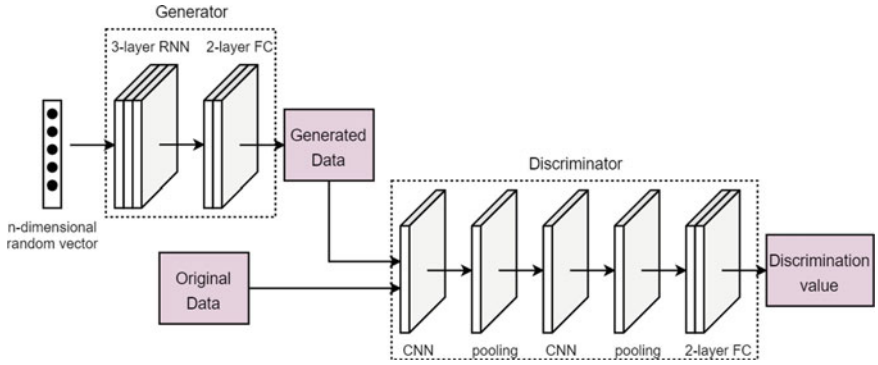


Fig. 4 Architecture of UIGAN

3n RNN nodes, where n represents the total number of component points in each frame of the layout data. The component points in the layout are obtained from the coordinate values of each vertex of the bounding box. It is observed that there are two fully connected layers, with each FC still containing 3n nodes.

The fully connected layer is used to further generate layout data. Through the fully connected layers, data with the same dimensions and time correlations with the original data can be generated. Through training, the sequence of the layout data that meets the needs can then be generated.

The discriminator is mainly composed of the CNN, pooling layer, and FC. Each CNN layer includes a 3×3 convolution kernel. The pooling layer adopts maximum pooling. The number of nodes in the FC of the first layer was set to 13, and the number of nodes in the FC of the second layer was set to one. The main task of the CNN is to extract the matrix's shape features. Because the layout data sequence is essentially the same as the image data, the trained CNN can effectively analyse the difference between the generated data and the original data. The three-layer FC's objective is to convert the extracted features into identification values to determine the data's authenticity.

3.4 Training of UIGAN

After preprocessing the dataset, it is also necessary to initialise the GAN's parameters. Generally, all parameters are either assigned a value of zero or conform to a normal distribution. After all initialisation work is completed, model training can be started.

The training process of UIGAN alternates between training the generator and training the discriminator [27]. In training the generator, one must ensure that the discriminator's parameters remain unchanged. In a similar vein, the process of training the discriminator requires the generator's parameters to be kept constant. This

process is an iteration, and each iteration includes two forward propagations and two backpropagations.

In training the generator, a piece of data is selected from the dataset (that is, layout data is extracted, which is generally stored in the form of a matrix). The matrix's number of rows is the number of layouts, set as m, and n represents the total number of three-dimensional coordinate dimensions of all the component points from the layout. Thus, the data is expressed as follows:

$$I_{mn} = F(O) \tag{1}$$

where O is the original dataset, and F is the layout data selected from O.

After selecting an action, one must generally obtain the number of layouts m and the number of spatial coordinates n of the component points. To ensure that the generated data are as similar to the original data as possible, it is necessary to set the generated data matrix's number of cycle iterations to m. Doing so ensures that the number of rows and columns of the generated data matrix are equal to those of the original layout data.

Next, the system needs to generate a random vector set as z, input the random vector into the generator, and, after m iterations, generate a matrix with the same dimensions as the original data R according to the algorithm described above, which is called the pseudo layout sequence Z. This is expressed by the following:

$$S = G(Z) \tag{2}$$

The generated layout sequence S is then sent to the discriminator to generate an eigenvalue T as follows:

$$T = D(S) \tag{3}$$

Finally, the generator parameters are adjusted according to the following error formula:

$$L_g(S) = -\sum \log(D(G(S))) \tag{4}$$

The adaptive moment estimation (Adam) technique is used to optimise all the parameters of the generator. In the process of training the identifier, it is ensured that the generator parameters remain unchanged. Similarly, a piece of data in the real action sequence is selected, such as in Eq. 1. After selecting an action, it is still necessary to obtain the number of layouts m and the number of spatial coordinates n of the component points. Similarly, the number of iterations of the generated loop is then set to m.

Through the method described above, the random variable z is obtained. After passing through the generator according to Eq. 2, the pseudo layout data matrix is still obtained. It is then input into the discriminator to obtain the characteristic value T1 according to the following equation:

$$T_1 = D(S) \tag{5}$$

In a process different from the one utilised for training the generator, the original data R is input into the discriminator to obtain the eigenvalue T_2 as follows:

$$T_2 = D(R) \tag{6}$$

Adjust the parameters of the appraiser according to the error formula:

$$L_d(S, R) = -\sum(\log(T_1) + \log(1 - T_2)) \tag{7}$$

Finally, Adam is used to optimise all the parameters of the discriminator. The two processes are iterated continuously. When the discriminator can no longer identify the data created by the generator, the training is deemed complete.

After the training, the generator needs to be extracted from the UIGAN system, the required dataset expansion tool. The continuous input of random variables can generate a large number of datasets to supplement contexts with an otherwise insufficient number of GUI datasets.

4 Experiment

The computer used in the experiment had 32 GB of DDR4 RAM and 8 GB of graphics memory through an NVIDIA GeForce RTX2080 graphics card accelerator.

Because the GUI dataset is stored as a basic digital sequence, a layout sequence can be abstracted into a matrix and then further abstracted into a vector through the transformation method. Therefore, the layout sequence can be transformed into a matrix or vector through abstract methods. When evaluating the similarity between two actions, one must only obtain the similarity of two matrices or vectors through mathematical methods.

At present, two evaluation methods were adopted:

The first method is Euclidean distance evaluation. This loss function is used for the similarity of matrices. Let the two matrices be A and B and the elements be a and b, respectively. The Euclidean distance can be expressed as follows:

$$L_e = \sum_{A,B} |a - b| \tag{8}$$

This method's main idea is to add the differences of numerical values to obtain the differences in the data and ignore the influence of other aspects. The advantages are intuitive and simple; the disadvantage is that significant calculation is necessary, especially for the square calculation of 1-norm.

The second method is cosine similarity evaluation. Cosine similarity measures the correlation between two vectors (x, y), which is defined as follows:

$$cos(x, y) = \frac{x \cdot y}{|x| \cdot |y|} \tag{9}$$

The calculation result is a decimal value from -1 to 1. Specifically, -1 means that the two vectors have opposite directions, 0 indicates that the two vectors are vertical, and 1 shows that the two vectors have the same direction.

This method describes the similarity of vectors in terms of vector direction. If the directions of the two vectors differ greatly, this will be reflected by this method. A combination of the two methods is used to measure the dataset generated by UIGAN from two angles and compare it with other semi-supervised and unsupervised algorithms.

Summarise the formula of cosine similarity and define the evaluation formula as follows:

$$E = \frac{L_e}{2 + cos(x, y)} \tag{10}$$

4.1 Experimental Results

Ideally, by the end of the UIGAN training, the generator and discriminator will have strong creation and identification abilities. The two form a state of confrontation: when the value of one loss function rises, the other will fall until the values balance.

The calculation results of the loss function are shown in Tables 1 and 2. By using Eq. 6, increasing the iterations results in the identification of the original data converging to one, indicating that the identification ability is increasing.

How to determine the loss result of the discriminator for the generator is indicated in Eq. 5. From this, it can be seen that the result generated by the discriminator for

Table 1 Calculation results of the Euclidean distance loss function

Number of iterations	Original data loss	Generated data loss
0	0.81	0.82
2000	0.97	0.84
4000	0.98	0.8
6000	0.99	0.85
8000	0.99	0.86
10,000	0.99	0.86
12,000	0.99	0.86
14,000	0.99	0.87

Table 2 Calculation results of discriminator loss and generator loss

Number of iterations	Discriminator loss	Generator loss
0	2.03	0.24
2000	2.07	0.22
4000	2.19	0.22
6000	2.34	0.21
8000	2.49	0.21
10,000	2.63	0.21
12,000	2.77	0.20
14,000	2.91	0.20

the generator gradually increases and approaches one. Stated differently, the data produced by the generator becomes increasingly similar to the original data.

The change of the loss function of the discriminator in the training process corresponds to Eq. 7, whilst the change of the loss function of the generator in the training process is represented by Eq. 4.

It can be seen that the two trends begin at opposite ends of a range and finally converge, which fully reflects their confrontation operation. The changes in the above parameters reflect the training process of UIGAN. Under the described hardware conditions, the total time consumed by the whole training process was measured at about seven hours.

The process of traditional GAN-that is, the GAN model if the RNN module is deleted-is as follows: input the original data into the traditional GAN, use the GAN to generate data of a certain scale, then calculate the evaluation value of the dataset (as in Eq. 9).

The experiment was carried out on the Rico dataset, and the results are shown in Table 3. Several layouts were randomly selected for the experiment. From the analysis of the Rico dataset experiment, it can be concluded that the similarity between the data generated by UIGAN and the original data was higher than that of the traditional GAN and the original data; that is, the RNN achieved the expected effect in the UIGAN system.

However, there exists a limited number of experiments and evaluation standards in related fields to serve as benchmarks in regards to judging whether the similarities between the dataset created by UIGAN and the original data were reasonable and well

Table 3 Experimental results of GUI generation based on the Rico dataset

Data object category	Traditional GAN (E value)	UIGAN (E value)
Text	2708.71	2693.35
Image	2723.54	2634.47
Button	2265.68	2116.81

generated. Nevertheless, the experimental results returned an evaluation value smaller than that of traditional GAN, showing that the algorithm improved upon the GAN in the field of GUI generation to a certain extent. This could be attributed to the fact that UIGAN can learn the dependencies between data sequences. The experimental results also show that UIGAN was able to generate similar GUI datasets.

5 Conclusion

The algorithm regularly deviates from the correct result of a GUI in generating a complex layout; thus, further improvements are necessary. On the whole, UIGAN performs well in experiments of dataset augmentation, going beyond initial expectations. Despite the lack of comparable experiments, the generated dataset meets the expected goal for GUI layout training and performs better than the traditional GAN based on the experimental results.

References

1. Kull A (2012) Automatic gui model generation: state of the art. In: 2012 IEEE 23rd international symposium on software reliability engineering workshops. IEEE, pp 207–212
2. Shneiderman B, Plaisant C, Cohen MS, Jacobs S, Elmqvist N, Diakopoulos N (2016) Designing the user interface: strategies for effective human-computer interaction, Pearson
3. Mistry M, Apte A, Ghodake V, Mane SB (2019) Machine learning based user interface generation. In: International conference on intelligent computing, information and control systems. Springer, Heidelberg, pp 453–460
4. Li Y, Yang Z, Guo Y, Chen X (2019) Humanoid: a deep learning-based approach to automated black-box android app testing. In: 2019 34th IEEE/ACM international conference on Automated Software Engineering (ASE). IEEE, pp 1070–1073
5. Wang W, Peng Y, Lu L, Lu Z, Bagheri M, Summers R (2017) Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: IEEE CVPR, vol 7
6. Simard P, Victorri B, LeCun Y, Denker JS (1991) Tangent prop-a formalism for specifying selected invariances in an adaptive network. In: NIPS, vol 91. Citeseer, pp 895–903
7. Simard PY, Steinkraus D, Platt JC et al (2003) Best practices for convolutional neural networks applied to visual document analysis. In: Icdar, vol 3
8. Holtzblatt K, Beyer H (2014) Contextual design: evolved. Synthesis Lect Human-Centered Inf 7(4):1–91
9. Deka B, Huang Z, Franzen C, Hibschman J, Afergan D, Li Y, Nichols J, Kumar R (2017) Rico: a mobile app dataset for building data-driven design applications. In: Proceedings of the 30th annual ACM symposium on user interface software and technology, pp 845–854
10. Jaderberg M, Simonyan K, Zisserman A et al (2015) Spatial transformer networks. In: Advances in neural information processing systems, pp 2017–2025
11. Puerta AR (1997) A model-based interface development environment. IEEE Softw 14(4):40–47
12. Elkoutbi M, Khriss I, Keller RK (2006) Automated prototyping of user interfaces based on uml scenarios. Autom Softw Eng 13(1):5–40
13. da Silva PP, Paton NW (2001) A uml-based design environment for interactive applications. In: Proceedings second international workshop on user interfaces in data intensive systems. UIDIS 2001. IEEE, pp 60–71

14. Bošnjak M, Rocktäschel T, Naradowsky J, Riedel S (2017) Programming with a differentiable forth interpreter. In: International conference on machine learning. PMLR, pp 547–556
15. Bunel RR, Desmaison A, Mudigonda PK, Kohli P, Torr P (2016) Adaptive neural compilation. Adv Neural Inf Process Syst 29:1444–1452
16. Balog M, Gaunt AL, Brockschmidt M, Nowozin S, Tarlow D (2016) Deepcoder: learning to write programs. arXiv preprint arXiv:1611.01989
17. Gaunt AL, Brockschmidt M, Singh R, Kushman N, Kohli P, Taylor J, Tarlow D (2016) Terpret: a probabilistic programming language for program induction. arXiv preprint arXiv:1608.04428
18. Ling W, Grefenstette E, Hermann KM, Kočiskỳ T, Senior A, Wang F, Blunsom P (2016) Latent predictor networks for code generation. arXiv preprint arXiv:1603.06744
19. Beltramelli T (2018) pix2code: generating code from a graphical user interface screenshot. In: Proceedings of the ACM SIGCHI symposium on engineering interactive computing systems, pp 1–6
20. Zhu Z, Xue Z, Yuan Z (2018) Automatic graphics program generation using attention-based hierarchical decoder. In: Asian conference on computer vision. Springer, Heidelberg, pp 181–196
21. Nguyen TA, Csallner C (2015) Reverse engineering mobile application user interfaces with remaui (t). In: 2015 30th IEEE/ACM international conference on Automated Software Engineering (ASE). IEEE, pp 248–259
22. Kevin M, Carlos B-C, Michael C, Richard B, Denys P (2018) Machine learning-based prototyping of graphical user interfaces for mobile apps. IEEE Trans Softw Eng 46(2):196–221
23. Karras T, Aila T, Laine S, Lehtinen J (2017) Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196
24. Aggarwal A, Mittal M, Battineni G (2021) Generative adversarial network: an overview of theory and applications. Int J Inf Manage Data Insights :100004
25. Girshick R (2015) Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 1440–1448
26. Razavian AS, Azizpour H, Sullivan J, Carlsson S (2014) Cnn features off-the-shelf: an astounding baseline for recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 806–813
27. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078

Robust Stabilization of Ball-Plate System with Higher-Order SMC

Suleiman U. Hussein, Mohammed B. Mu'azu, Sikiru T. Humble, Chichebe M. Akachukwu, Eseoghene Ovie, and Umar M. Mustapha

Abstract The work present improved methods for stabilizing ball-plate system with different variant of sliding mode control (SMC) like classical SMC, SSMC, RENN-optimized SMC, and SSTSMC based on their level of robustness in the presence of bounded disturbances. In this research, double-feedback loop was used in controlling the system, where PID controller, tuned with wAFSA, was used for the inner loop, while the variant SMC was used for the outer loop control. The results from the variant SMCs were compared as regards performance (rise-time and steady-state error) and disturbance rejection in order to determine the best variant. Therefore, best variant was the smoothed super-twisting SMC (3.2462 s and 1.2244e-04%, respectively) and robust disturbance rejection ability. The work is based on simulation in MATLAB.

Keywords Ball-plate system · SMC · Chattering · RENN · Disturbance rejection

1 Introduction

Stabilization is one of the major goals in most control system design; therefore, the importance of this goal cannot be overemphasized. In view of this, stabilizing the ball precisely at a predefined point is a very critical challenge. Furthermore, when the ball is required to be held at a specific position, it is referred to as stabilization control, while when it is to follow a predefined path, it is regarded as trajectory tracking. Therefore, these goals are challenging in view of high precision [1].

The ball-plate system is used as a test bench for nonlinear dynamic problems like trajectory planning and following in control design [2]. This model is usually employed in testing different control algorithms [3], due to its properties like under-actuation, instability, and nonlinearity [4].

This system can be further described as a system with an unconstrained object that moves freely on a platform without ability of environment recognition. Therefore, this object does not have the ability of controlling itself [5]. Also, the ball's velocity

S. U. Hussein (✉) · M. B. Mu'azu · S. T. Humble · C. M. Akachukwu · E. Ovie · U. M. Mustapha
Ahmadu Bello University, Zaria, Nigeria
e-mail: elsuligh@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_60

on the plate could be high; therefore, designing a controller that will suit this system could pose a serious challenge [6].

The ball-plate system could be applied when designing controllers in areas like unmanned ground vehicles, rocketry, robotics, unmanned aerial vehicles, and attitude control of satellites [6–9]. As a result, stabilization of this system is of utmost importance, especially in the presence of disturbances.

Therefore, [6] proposed a novel dual proportional-derivative controller for controlling the ball-plate system. The work considered a dual PD controller with a compensator that is nonlinear which was added with the aim of guaranteeing stability of the system in closed loop. The proposed controller gave a better result when compared with fuzzy logic controller in terms of stabilization; however, the ability for the system to reject bounded disturbances was not considered in the study. Also, [10] considered proportional-derivative controllers in cascaded form for trajectory control of ball-plate system where particle swarm optimization was used for the tuning of PD controller parameters in both loops. The results showed successful tracking however, when angular frequency of the trajectory increases the tracking speed also increase, which led to increase in the occurrence of steady-state error. Furthermore, [11] focused on addressing chattering in fuzzy SMC for position and path following in ball-plate. In the work, introduction of sliding mode controller was made in the outer loop of the double-feedback loop. Also, the inner loop was achieved with linear algebraic method. Furthermore, fuzzy logic with genetic algorithm for optimal determination of the fuzzy parameters was employed with the essence of minimizing chattering in the SMC. The simulation result showed good stabilization and trajectory tracking system; however, disturbance rejection ability of the controller was not considered. In the work of [12] where time delay was considered in the ball-plate system using SMC with respect to model reference, the results showed fast response from the controller in position and trajectory tracking; however, introduction of disturbances to test the robustness of the controller was not considered. The study in [13] proposed approximate model for stabilizing ball-plate system, where controlled Lagrangian method was used to stabilize the system by deriving the control rule using the approximate solution from the matching condition. The results from the study showed that the ball on the plate was stabilized; however, the effect of disturbances on the system was not investigated. In [14] which proposes various methods for minimizing chattering in SMC with emphasis on the ball-plate system, the work first of all used smoothed SMC (SSMC) for reducing chattering, after which recurrent Elman neural network-optimized SSMC (RENNOSSMC) was developed for the system for further minimization in chattering. Also, developed for the system is smoothed super-twisting SMC (SSTSMC) for further reduction or possible elimination of chattering. The results of the RENNOSSSMC and SSTSMC show a very good reduction in chattering; however, the work did not consider stabilization of the system as well as effect of disturbances on it. In addition, [15] proposed SMC for implementation of ball-plate with sigmoid function selected for chattering reduction in the controller. The results showed that position and trajectory tracking were achieved; however, the

effect of disturbances on the system was not considered. Furthermore, [16] investigated the stabilization of the ball-plate using feedback linearization to attain stability; however, the effect of disturbance rejection was not considered.

Due to the effect of chattering in classical SMC, this work is an extension of the work of [14] where different methods for reducing chattering in SMC but stabilizing the ball-plate with the developed controllers was not considered. Therefore, this work focuses on stabilization of the system with respect to bounded disturbances using the developed controllers by considering robustness of the controllers in terms of rise time, steady-state error, percentage overshoot, and disturbance rejection. Results from the variant sliding mode controllers would be compared, and the best variant would be chosen for improved stabilization and disturbance rejection for the system.

The remainder of this work is ordered as follows: Sect. 2 talks about ball-plate model, while 3 gives variant SMC design. Section 4 contains results with discussions. Lastly, Sects. 5 and 6 cover the conclusion and future works.

2 The Ball-Plate Model

Basic components of this system include ball, flat square plate, camera, and two of DC stepper motors. Where function of the camera is to sense the point that the ball is, while the two DC motors perform the action of tilting the plate to keep the ball in the referenced position. Figure 1 is schematic of the system.

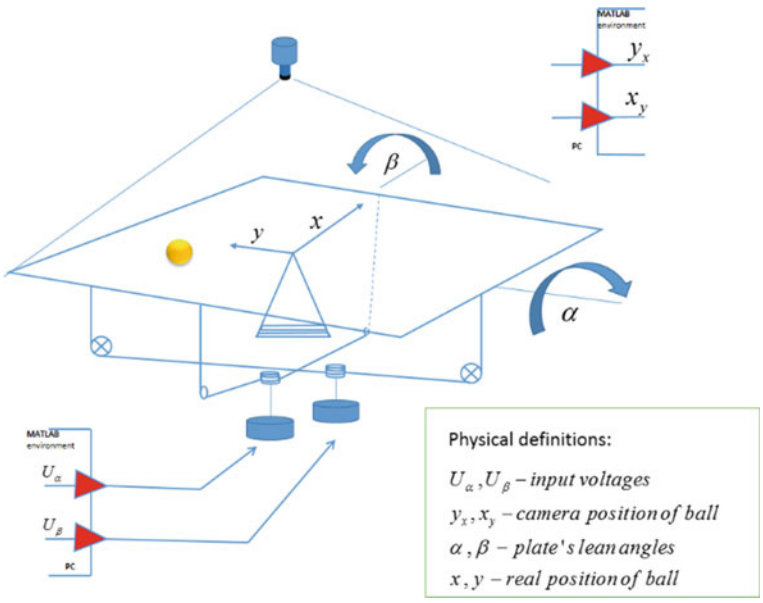


Fig. 1 Ball-plate schematic [17]

Dynamic model of the system was achieved with Euler–Lagrange principles. Given the generalized Euler–Lagrange representation of the system in Eq. (1)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L(q_i,\dot{q}_i,t)}{\partial \dot{q}_i}\right)-\frac{\partial L(q_i,\dot{q}_i,t)}{\partial q_i}=F_{qi},\quad 1\le i\le n \tag{1}$$

L stands for system's Lagrangian, given as

$$L\left(q_i,\dot{q}_i,t\right)=T(\dot{q}_i,t)-V(q_i,t) \tag{2}$$

$T(\dot{q}_i,t)$ and $V(q_i,t)$ are the kinetic and potential energy, respectively, while q_i stands for generalized coordinates and F_{q_i} represents generalized forces.

The coordinates of the system are as follows:

$$q=\begin{bmatrix}x\\y\\\alpha\\\beta\end{bmatrix}=\begin{bmatrix}x\\y\\\theta_x\\\theta_y\end{bmatrix}$$

System's kinetic energy:

$$T=\frac{1}{2}\left\{\left(J_{p_x}\dot{\theta}_x^2+J_{p_y}\dot{\theta}_y^2\right)+\left(m+\frac{J_B}{R^2}\right)\left(\dot{x}^2+\dot{y}^2\right)+J_B\left(\dot{\theta}_x^2+\dot{\theta}_y^2\right)+m\left(x\dot{\theta}_x+y\dot{\theta}_y\right)^2\right\} \tag{3}$$

System's potential energy:

$$V=\mathrm{mgy}\sin\theta_y+\mathrm{mgx}\sin\theta_x \tag{4}$$

Equation (2) becomes,

$$L=\frac{1}{2}\left\{\left(J_{p_x}\dot{\theta}_x^2+J_{p_y}\dot{\theta}_y^2\right)+\left(m+\frac{J_B}{R^2}\right)\left(\dot{x}^2+\dot{y}^2\right)+J_B\left(\dot{\theta}_x^2+\dot{\theta}_y^2\right)+m\left(x\dot{\theta}_x+y\dot{\theta}_y\right)^2\right\}-\left(\mathrm{mgx}\sin\theta_x+\mathrm{mgy}\sin\theta_y\right) \tag{5}$$

Equation (5) is substituted in (1) to give:

$$\left(m+\frac{J_B}{R^2}\right)\ddot{x}-\mathrm{mx}\left(\dot{\theta}_x\right)^2+\mathrm{mg}\sin\theta_x-\mathrm{my}\dot{\theta}_x\dot{\theta}_y=0 \tag{6}$$

$$\left(m+\frac{J_B}{R^2}\right)\ddot{y}-\mathrm{my}\left(\dot{\theta}_y\right)^2+\mathrm{mg}\sin\theta_y-\mathrm{mx}\dot{\theta}_x\dot{\theta}_y=0 \tag{7}$$

$$\left(\mathrm{mx}^2 + J_B + J_{P_x}\right)\ddot{\theta}_x + \mathrm{mxy}\ddot{\theta}_y + \mathrm{mgx}\cos\theta_x + 2\mathrm{mx\dot{x}}\dot{\theta}_x + \mathrm{mx\dot{y}}\dot{\theta}_y + \mathrm{my\dot{x}}\dot{\theta}_y = \Gamma_x \tag{8}$$

$$\left(\mathrm{my}^2 + J_B + J_{\mathrm{Py}}\right)\ddot{\theta}_y + \mathrm{mxy}\ddot{\theta}_x + \mathrm{mgy}\cos\theta_y + 2\mathrm{my\dot{y}}\dot{\theta}_y + \mathrm{mx\dot{y}}\dot{\theta}_x + \mathrm{my\dot{x}}\dot{\theta}_x = \Gamma_y \tag{9}$$

Resolving Eqs. (6), (7) into the equivalent second-order dynamic equations gives:

$$\ddot{x} = 1/J_{\mathrm{BM}}(\mathrm{mx}(\dot{\theta}_x)^2 + \mathrm{my}\dot{\theta}_x\dot{\theta}_y - \mathrm{mg}\sin\theta_x)$$

$$\ddot{y} = 1/J_{\mathrm{BM}}(\mathrm{my}(\dot{\theta}_y)^2 + \mathrm{mx}\dot{\theta}_x\dot{\theta}_y - \mathrm{mg}\sin\theta_y)$$

$$\ddot{\theta}_x = 1/J_{\mathrm{BMP}x}(-2\mathrm{mx\dot{x}}\dot{\theta}_x - \mathrm{my\dot{x}}\dot{\theta}_y - \mathrm{mx\dot{y}}\dot{\theta}_y - \mathrm{mxy}\ddot{\theta}_y - \mathrm{mgx}\cos\theta_x + \Gamma_x)$$

$$\ddot{\theta}_y = 1/J_{\mathrm{BMP}x}(-2\mathrm{my\dot{y}}\dot{\theta}_y - \mathrm{my\dot{x}}\dot{\theta}_x - \mathrm{mx\dot{y}}\dot{\theta}_x - \mathrm{mxy}\ddot{\theta}_x - \mathrm{mgx}\cos\theta_y + \Gamma_y)$$

where the denominator terms are as follows:

$$J_{\mathrm{BM}} = m + J_B/R^2;\ J_{\mathrm{BMP}x} = \mathrm{mx}^2 + J_B + J_{\mathrm{Px}};\ J_{\mathrm{BMP}y} = my^2 + J_B + J_{\mathrm{Py}}$$

Equations (6), (7) describe the ball's movement, while Eqs. (8), (9) described ball's position and speed. The plate is assumed to stay horizontally in steady state, which means both angles of inclination would be zero. Assuming the angles does not change beyond therefore, the sine function would be replaced with its argument [17, 18].

Inertia of the ball is computed as

$$J_B = \frac{5}{2}\mathrm{mR}^2 \tag{10}$$

Therefore, the system's mathematical equation is taken as

$$\ddot{x} = 0.7143g\theta_x \tag{11}$$

$$\ddot{y} = 0.7143g\theta_y \tag{12}$$

The DC stepper motor that is responsible for tilting the plate has a transfer function given as [19]:

$$\mathrm{TF} = k_m/(\tau_m s + 1) \tag{13}$$

k_m represents static gain of the servo system and represents time constant of servo system. Values for motor parameters [17, 18]:

k_m is 0.1878 rad/s, while τ_m takes 0.187 s. Resulting to

$$\mathrm{TF} = \frac{1}{0.996s + 5.325} \tag{14}$$

3 Controller Design

The process of developing a controller to achieve some certain desirable dynamics of a system is termed controller design. Double feedback loop is adopted for the controller design where proportional–integral–derivative (PID) controller tuned with weighted artificial fish swarm algorithm that controls the stepper motor is used for the inner-loop control as in [17], while the ball's position is controlled by the SMC in outer loop. The outer loop was first controlled by classical SMC, after which it was replaced by smoothed SMC (SSMC), then recurrent Elman neural network-optimized SSMC (RENNOSSMC) was used to replace the later as in [14], and then finally, smoothed super-twisted SMC (SSTSMC) was further used to replace the RENNOSSMC for the purpose of choosing best type of SMC for stabilizing the system. Figure 2 is block representation of the controlled ball-plate system.

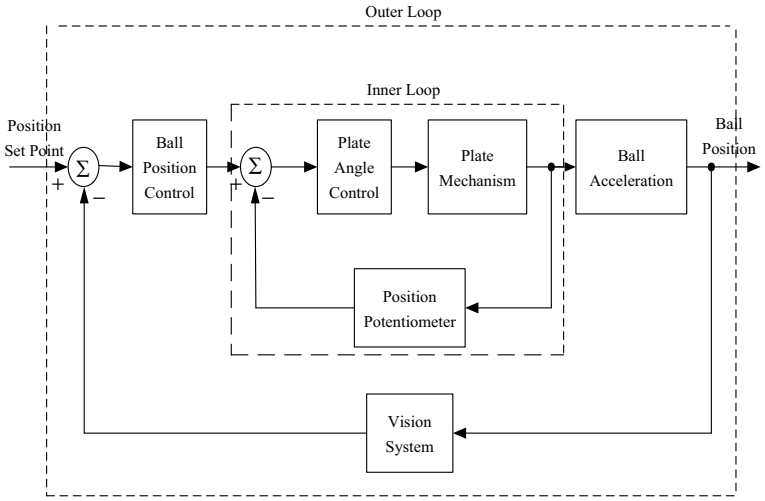


Fig. 2 Block diagram for controlled ball-plate system

3.1 Classical SMC Design for the Ball-Plate System

To design the sliding surface, let

$$\dot{s}(x,t) = 0 \tag{15}$$

$$s(x,t) = \left(\frac{d}{\mathrm{d}t} + \lambda\right)^{n-1} e \tag{16}$$

λ is strictly a non-negative constant and is the system's order. The system while in sliding mode is a function of λ. Sliding surface from Eq. (11):

$$s = \dot{e} + \lambda e \tag{17}$$

$e = x_r - x$ represents tracking error, and desired position while is the actual position. Making reference to Eqs. (15), (16) becomes

$$\dot{s} = \ddot{e} + \lambda \dot{e} \tag{18}$$

From Eq. (11), let $b = \frac{5g}{7}$ and $\theta_x = \alpha$, and substituting these into (18) with

$$\dot{s} = 0$$

Equivalent control of the system becomes

$$u_{eq} = b^{-1}\{\ddot{x}_r + \lambda \dot{x}_r - \lambda \dot{x}\} \tag{19}$$

Once the state variable reaches the sliding surface, Eq. (19) is responsible for keeping it onto that surface.

Since sliding mode control law can be defined as

$$u_{\mathrm{smc}} = u_{\mathrm{eq}} + u_n \tag{20}$$

where $u_n = k\mathrm{sgn}(s)$ is the switching control that is responsible for switching the state variable onto the sliding surface in finite time. Therefore, Eq. (20) becomes

$$u = b^{-1}\{\ddot{x}_r + \lambda(\dot{x}_r - \dot{x}) + k\mathrm{sgn}(s)\} \tag{21}$$

However, it is imperative to ascertain whether the state of the system will converge on the sliding surface when the control law is applied. Therefore, $s \to 0$ at $t \geq 0$ by applying Lyapunov's method of stability.

Thus, assuming a function $v(x,t)$ of the state x over time t such that:

$$v(x,t) = \frac{1}{2}\{s(x,t)\}^2 \tag{22}$$

$$\dot{v}(x,t) = \frac{\partial}{\partial t}\left\{\frac{1}{2}[s(x,t)]^2\right\} \tag{23}$$

$$\dot{v}(x,t) = \dot{s}s \tag{24}$$

Substituting Eq. (18) into (24), we have

$$\dot{v}(x,t) = (-bu + \ddot{x}_r + \lambda\dot{e})s \tag{25}$$

and putting Eq. (21) into (25) gives

$$\dot{v}(x,t) = \left\{-b\left[b^{-1}(\ddot{x}_r + \lambda\dot{e}) + k\text{gn}(s)\right] + \ddot{x}_r + \lambda\dot{e}\right\}s \tag{26}$$

But $|s| = \text{sgn}(s)s$, therefore,

$$\dot{v}(x,t) = -k|s| \tag{27}$$

$$\dot{v}(x,t) < 0 \tag{28}$$

Therefore, $\dot{v}(x,t)$ is negative definite since k is positive definite, which implies that the likelihood of the state of the system converging on the sliding surface is guaranteed when the controller is applied to the system.

3.2 SSMC Design for the Ball-Plate System

The SSMC is gotten from Eq. (21) using saturation function to replace the signum function which gives Eq. (29)

$$u = b^{-1}\{\ddot{x}_r + \lambda(\dot{x}_r - \dot{x}) + k\text{sat}(s)\} \tag{29}$$

3.3 RENNOSSMC Design for the Ball-Plate System

The RENNOSSMC is as a result of tuning the k in Eq. (29) with recurrent Elman neural network to get the optimal value, as in [14].

3.4 SSTSMC Design for the Ball-Plate System

The algorithm is given by Eq. (30) as [14]:

$$u_n = u_{\text{na}} + \dot{u}_{\text{nb}} \tag{30}$$

Substituting Eq. (30) into (20) gives the SSTSMC where

$$u_{\text{na}} = -\sqrt{k} \times \sqrt{|s|} \times \text{sat}(s) \tag{31}$$

$$\dot{u}_{\text{nb}} = -1.1 \times k \times \text{sat}(s) \tag{32}$$

4 Results and Analysis

In this section, responses of the system to the different SMCs as regard stabilization and disturbance rejection are discussed.

4.1 System Response with Classical SMC

With a unit step input to the system, while SMC ($k = 0.5$), Fig. 3 shows that the ball is stabilized with respect to the reference. More so, Table 1 gives more detail of the system response. Furthermore, when disturbance of 0.04 step was introduced

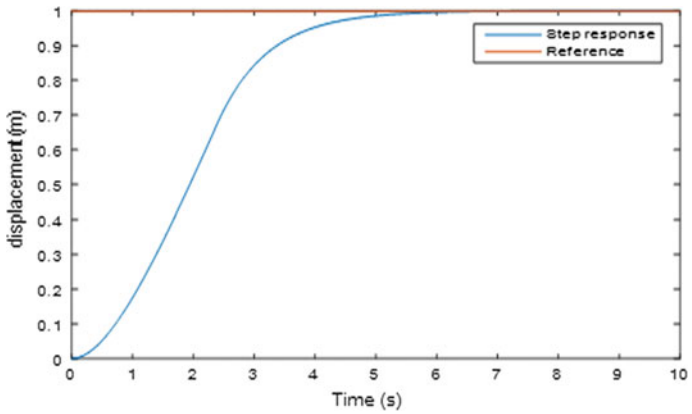


Fig. 3 System's unit step response

Table 1 Performance characteristic of B&P with SMC

Characteristics	Value
Rise time	2.6534 s
Settling time	4.7149 s
Overshoot	0%
Peak	1.0000 m
Peak time	10.0000 s
Steady-state error	0

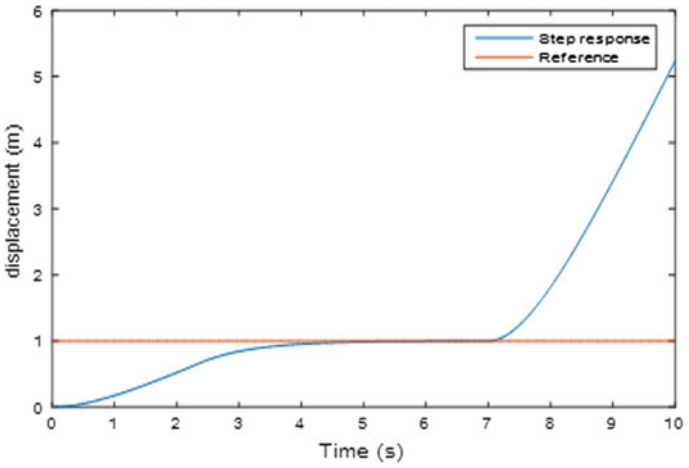


Fig. 4 Unstable system's response with 0.4 step disturbance

based on trial and error at 7 s, the system response remained stable which means the controller is robust. However, when the magnitude of the disturbance increased to 0.4 step, it can be seen that the system became unstable as shown in Fig. 4 and the steady-state error $\rightarrow \infty$.

The gain (k) of the SMC was then increased to 200, and it was observed that the system was stable with respect to 0.4 step disturbance. This is, however, at the cost of a relatively higher chattering in the effort of the SMC.

4.2 System Response with SSMC

With a unit step input to the system, while SSMC ($k = 5$), Fig. 5 shows that the ball is stabilized with respect to the input. Furthermore, when with the introduction of 0.04 step disturbance based on trial and error at 7 s, the system response remains the same which means the controller is robust; however, with the increment of the disturbance to 0.4 step as shown in Fig. 6, the system response was still stable but

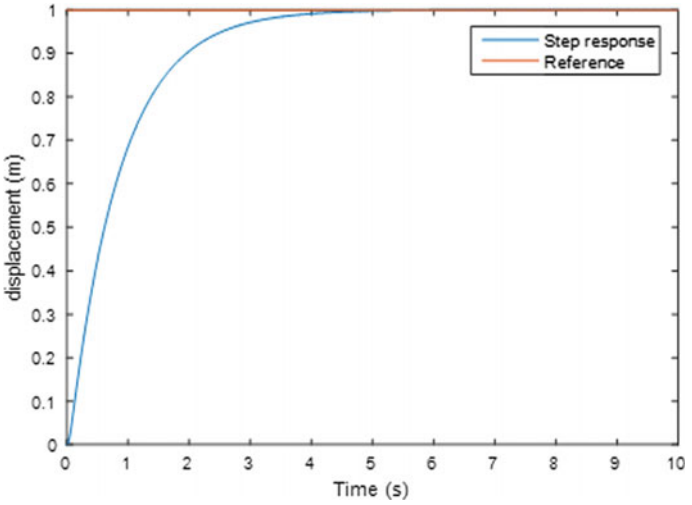


Fig. 5 System's unit step output (SSMC)

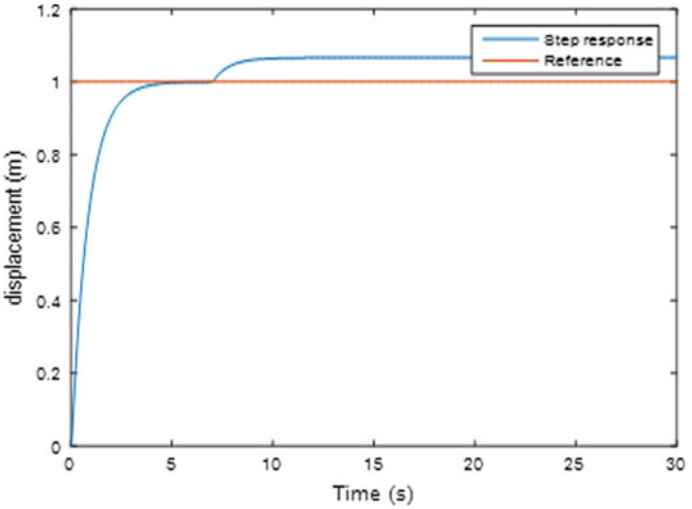


Fig. 6 System's response with step disturbance 0.4 (SSMC)

with steady-state error of 0.0667 m (66.7 mm). Table 2 gives a detailed characteristic of the response.

Table 2 Performance characteristic with SSMC

Characteristics	Value
Rise time	1.8292 s
Settling time	3.3111 s
Overshoot	0%
Peak	1.0000 m
Peak time	10.0000 s
Steady-state error	0

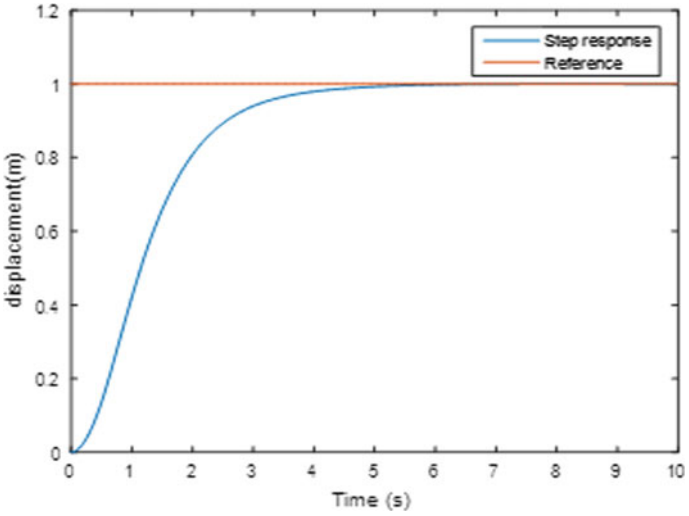


Fig. 7 System's unit step response (RENNOSSMC)

4.3 System Response with RENNOSSMC

With unit step reference to the system, Fig. 7 shows that the ball stabilizes with respect to the input as shown. Furthermore, with the introduction of 0.04 step disturbance at 7 s, the system remains stable but with steady-state error of 0.0877 m (87.7 mm) which can be seen in Fig. 8. However, Fig. 9 shows that with increased disturbance to 0.4 step, it became unstable with steady-state error $\rightarrow \infty$ as seen in Fig. 9. Table 3 gives the performance characteristics of the system.

4.4 System Response with SSTSMC

With a unit step input to the system, the ball was stabilized as seen in Fig. 10. Furthermore, with the introduction of 0.04 step disturbance at 7 s, the system was

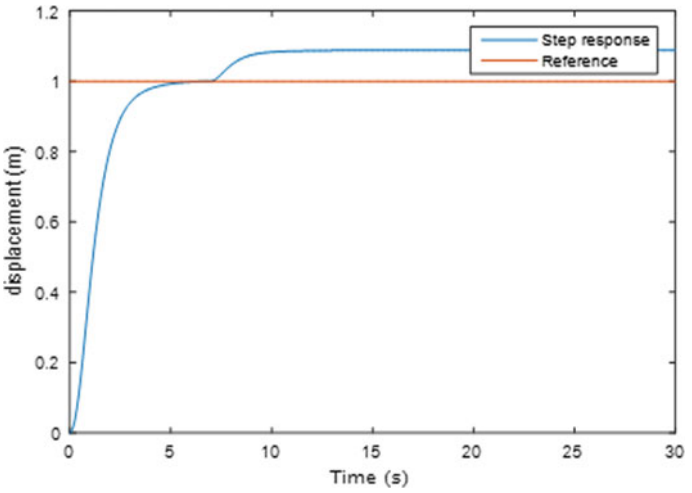


Fig. 8 System response with disturbance of 0.04 step (RENNOSSMC)

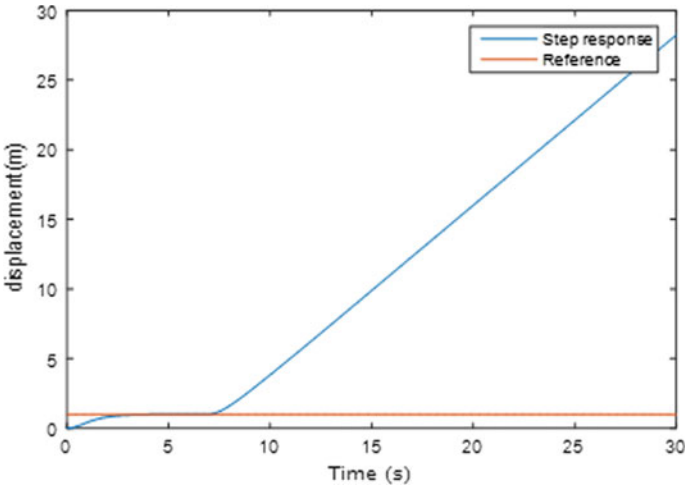


Fig. 9 System response with 0.04 step disturbance (RENNOSSMC)

Table 3 Performance characteristics of the system with RENNOSSMC

Characteristics	Value
Rise time	2.1374 s
Settling time	4.0469 s
Overshoot	0%
Peak	1.0004 m
Peak time	8.0734 s
Steady-state error	0.0004 m

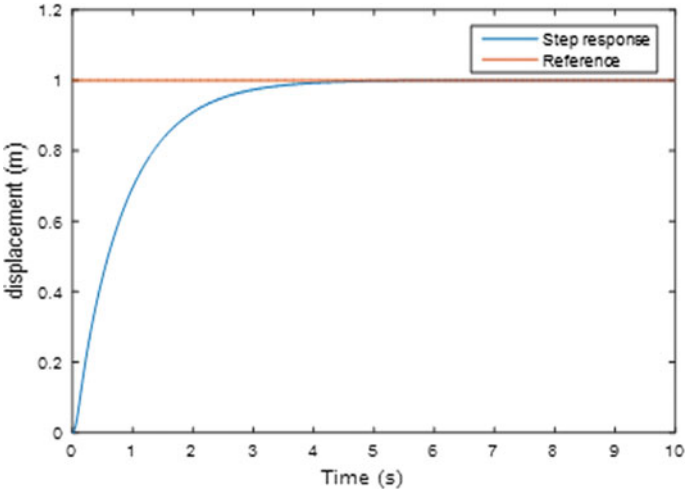


Fig. 10 Response of system with unit step input (SSTSMC)

stable without steady-state error, and with a further increment of the disturbance (0.4 step), it remains stable without steady-state error. It, however, experienced the effect of disturbance with magnitude 0.0063 m, and this lasted for 2.5 s as seen in Fig. 11. With disturbance was moving to 10 step, the system remained stable without steady-state error but with 2.7724 m as the magnitude of the disturbance effect, and this stayed for 4.5 s which can be seen in Fig. 12. This is an evidence that the system

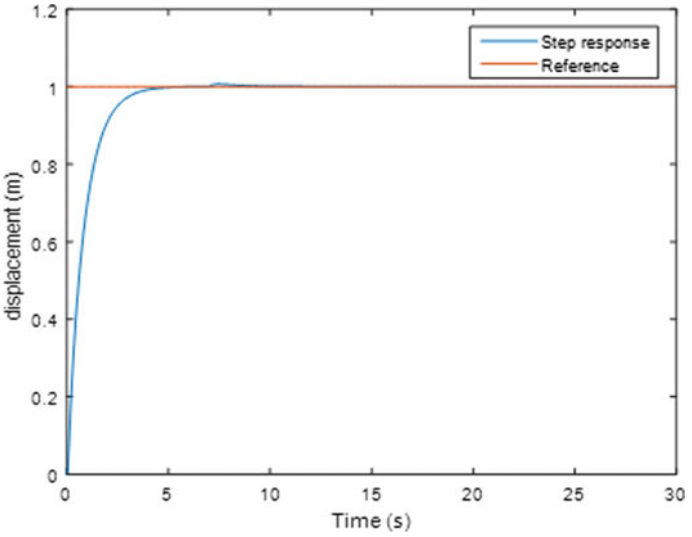


Fig. 11 System response with 0.4 step disturbance (SSTSMC)

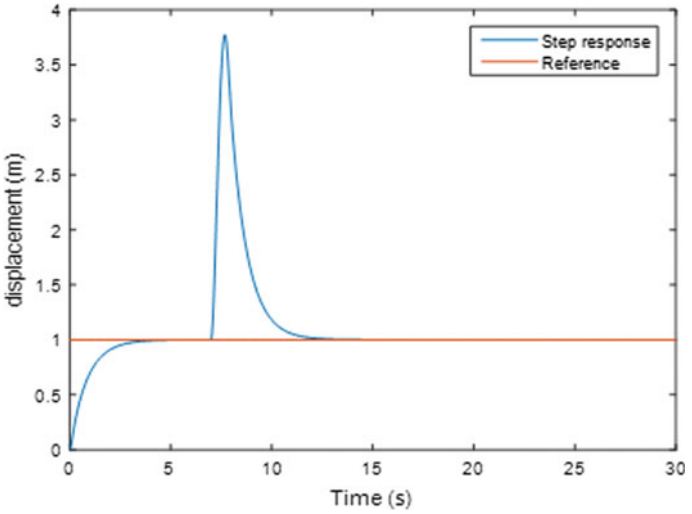


Fig. 12 System response with 10 step disturbance (SSTSMC)

Table 4 Performance characteristics of the system with SSTSMC

Characteristics	Value
Rise time	1.7917 s
Settling time	3.2462 s
Overshoot	1.2244e–04%
Peak	1.0000 m
Peak time	9.4636 s
Steady-state error	0

with SSTSMC is highly relatively robust due to it disturbance rejection capability. Table 4 gives output characteristic of the plant.

4.5 Comparative Analysis of the Different SMC Schemes

Comparison between the different sliding mode control schemes in terms of stabilization of the ball and responses of the controllers with unit step input is shown in Fig. 13. Also, comparison between the different sliding mode control schemes was done in Table 5. Looking at the overall performance of each controller, it shows that the SSTSMC outperformed the other controllers because it has very fast response time, excellent robustness, and no steady-state error.

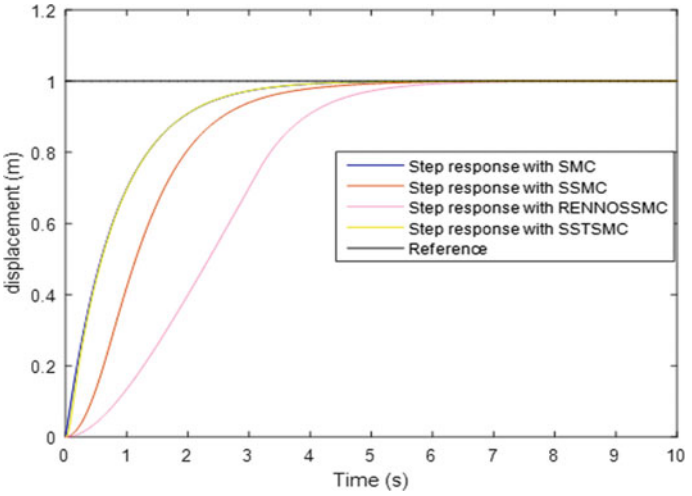


Fig. 13 Step responses of the ball-plate system for all the different SMCs

Table 5 Comparison of the different SMC schemes

Controller	Response time	Steady-state error
SMC	Fast	None
SSMC	Very fast	None
RENNOSSMC	Fast	Exist
SSTSMC	Very fast	None

5 Conclusion

In this work, ball stabilization has been achieved with all designed SMCs. However, the SSTSMC showed a better robust stabilization and disturbance rejection ability for the ball on the plate over the classical SMC, SSMC, and RENNOSSMC. Therefore, the SSTSMC is better for robust stabilization of ball-plate system.

6 Future Work

Trajectory tracking of ball-plate system with the different designed SMCs will be considered, and comparison will be done to determine the best in terms of robustness and tracking error.

References

1. Hongrui W, Yantao T, Siyan F, Zhen S (2008) Nonlinear control for output regulation of ball and plate system. In: Control conference. 27th Chinese, 2008. CCC, pp 382–387
2. Duan H, Tian Y, Wang G (2009)Trajectory tracking control of ball and plate system based on auto-disturbance rejection controller. In: Asian control conference, 2009. ASCC 2009, pp 471–476
3. Dong X, Zhang Z, Chen C (2009) Applying genetic algorithm to on-line updated PID neural network controllers for ball and plate system. In: Innovative computing, information and control (ICICIC), 2009 fourth international conference on, 2009, pp 751–755
4. Ho M-T, Rizal Y, Chu L-M (2013) Visual servoing tracking control of a ball and plate system: design, implementation and experimental validation. Int J Adv Robot Syst 10
5. Dong X, Zhao Y, Xu Y, Zhang Z (2011) Design of PSO fuzzy neural network control for ball and plate system. Int J Innov Comput Inf Control 7:7091–7103
6. Sergio G-C, Marco M-A, Rubio JDJ, Ortíz-Rodriguez F, Yu W, Aguilar-Ibáñez C (2014) Dual PD control regulation with nonlinear compensation for a ball and plate system. Math Prob Eng 2014
7. Mukherjee R, Minor MA, Pukrushpan JT (2002) Motion planning for a spherical mobile robot: revisiting the classical ball-plate problem. J Dyn Syst Meas Contr 124:502–511
8. Oriolo G, Vendittelli M (2005) A framework for the stabilization of general nonholonomic systems with an application to the plate-ball mechanism. Robot IEEE Trans 21:162–175
9. Negash AH (2011)Design of fuzzy sliding mode controller for the ball and plate system. M.Sc, Electrical and Computer Engineering, Addis Ababa University, Addis Ababa
10. Roy P, Kar B, Hussain I (2015)Trajectory control of a ball in a ball and plate system using cascaded PD controllers tuned by PSO. In: Proceedings of fourth international conference on soft computing for problem solving, pp 53–65
11. Andinet N, Singh PN (2015) Position control and tracking of ball and plate system using fuzzy sliding mode controller. Afro-Euro Conf Ind Adv 2015:123–132
12. Morales L, Camacho O, Leica P, Chavez D (2017) Sliding-mode control based on a model reference applied to a non-linear ball and plate system with time delay. In: 2017 IEEE 3rd Colombian conference on automatic control (CCAC), 2017, pp 1–6
13. Yıldız HA, Gören-Sümer L (2017) Stabilizing of ball and plate system using an approximate model. IFAC-PapersOnLine 50:9601–9606
14. Suleiman HU, Mu'azu MB, Zarma TA, Salawudeen AT, Thomas S, Galadima AA (2013) Methods of chattering reduction in sliding mode control: a case study of ball and plate system. In: 2018 IEEE 7th international conference on adaptive science & technology (ICAST), pp 1–8
15. Bang H, Lee YS (2018) Implementation of a ball and plate control system using sliding mode control. IEEE Access 6:32401–32408
16. Huang W, Zhao Y, Ye Y, Xie W (2019) State feedback control for stabilization of the ball and plate system. In: 2019 Chinese control conference (CCC), pp 687–690
17. Suleiman HU, Muhammed B, Sikiru T, Umoh I, Salawudeen A (2017) Trajectory tracking control of ball on plate system using weighted artificial fish swarm algorithm based PID. In: Electro-technology for national development (NIGERCON), 2017 IEEE 3rd international conference, pp 554–561
18. HUMUSOFT (2012) Ball and plate apparatus. In: User's manual, ed. Praque: HUMUSOFT, p 33
19. Chen C (1993) Analog and digital control system design: transfer-function. In: State-space, and algebraic methods. Saunders College, p 551

Prototype Machine for Traditional and Technological Ophthalmic Tests, Using Convergence Analysis

Sonia Cárdenas-Delgado, Mauricio Loachamin-Valencia, Paulette Parra Suárez, and Steeven Taipicaña Cayambe

Abstract Vision is one of the human senses that influences the way we perceive the world. There are traditional and technological ophthalmic tests used to identify visual deficiencies. Two visual tests using virtual reality were implemented in our project. The results of both studies showed correlation with the traditional tests. To improve the accuracy of all tests, we proposed to build an ophthalmic testing machine that allows traditional and technological tests to be integrated. This paper presents the prototype design of an ophthalmic tests machine that will allow to integrate different devices and tests. The general method of development of Ulrich products was used for the design stages of the machine. An adaptive mesh analysis was used to verify the chosen materials and components with more forces applied on them. This obtained a smaller divergence of 2.5% on all elements.

Keywords Ophthalmic assessment · Prototype · Machine · Design · Visual test

1 Introduction

Vision is one of the human senses that influences the way we perceive the world. To pilot aircraft, all visual abilities like dynamic visual acuity, wide visual field, depth vision, color vision, and rapid adaptation to light and darkness is required [4, 13].

S. Cárdenas-Delgado (✉) · M. Loachamin-Valencia · P. P. Suárez · S. T. Cayambe
Departamento de Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Av. General Rumiñahui s/n, 171-5-231B Sangolquí, Ecuador
e-mail: secardenas@espe.edu.ec

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

P. P. Suárez
e-mail: jpparra4@espe.edu.ec

S. T. Cayambe
e-mail: sataipicana@espe.edu.ec

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_61

Depth perception is of the utmost importance, as many military aviation maneuvres rely on accurate perception of distance [9, 13]. Stereoscopic vision is key for performance of pilots in aeronautics according to the study. This study demonstrated several cases in which the lack of this vision led to accidents. Most of them occurs during the cadet pilots' training [13]. This shows that stereoscopic vision is necessary to assure safety throughout the flight. It also required in order to improve aerial maneuvres.

Within this context, color vision is the visual ability to differentiate between certain colors [14]. The deficiency of color vision impacts the performance in daily life activities or jobs in which chromatic discrimination ability is required. Studies related to this topic show the importance of this ability in pilots to see through displays. This is needed in any kind of flight [10, 12].

There are several traditional tests to identify these visual deficiencies. Some of these tests are LANG test [6, 7], Ishihara test [3], Titmus test [8].

As more modern, higher-performing aircrafts are integrated in the Armed Forces, there is a higher demand to assess the different human factors such as psychomotor coordination, visual abilities, physiological performance, and cognitive and perceptual abilities from pilots [5].

To improve the assessment of pilot candidates visual abilities, computer-based testing needs to be applied. Two visual tests using virtual reality were implemented in our project in [1] and [2]. Both of these studies assessed stereopsis and vision colour through experimental studies using virtual reality.

The results of both studies showed correlation with the traditional tests. However, to improve the accuracy of the results, we proposed to build an ophthalmic testing machine that allows traditional and technological tests to be integrated.

This paper presents the design prototype of an ophthalmic testing machine that allows integrating traditional and technological tests. The general method of development of Ulrich products was used for the design stages of the machine.

2 Method

Although it is true, there are several ways to develop a product, depending on the application or the sector to which said product is directed, this work uses a general product development process detailed by [12], this process is made up of the following stages that will be described below (Fig. 1a).

2.1 Planning

This phase or stage includes the idea of the product to be materialized and the elaboration of a formal development proposal [12]. The idea for the creation of this

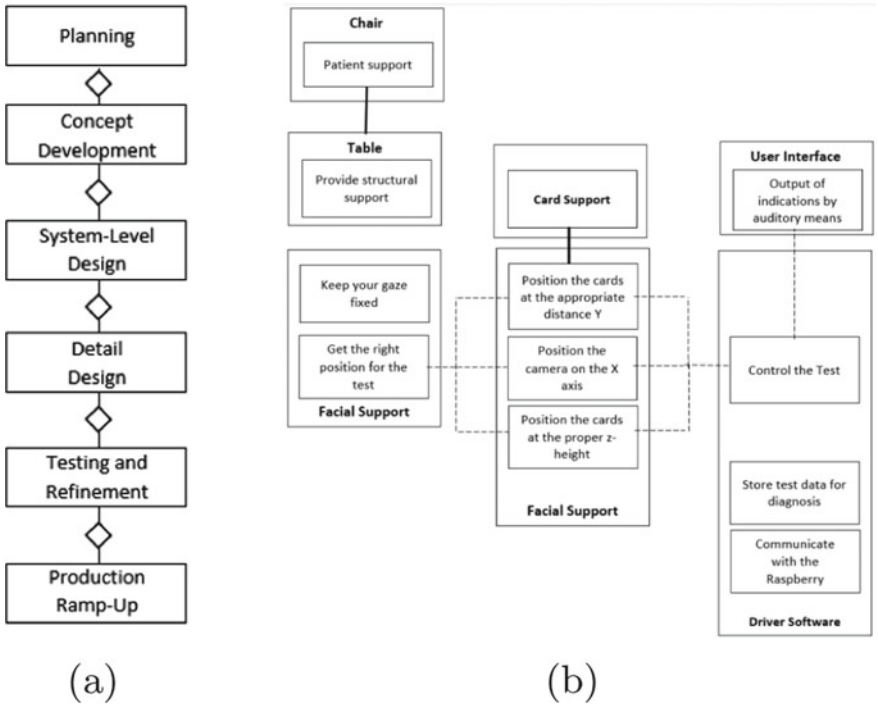


Fig. 1 **a** General process for product development **b** product online

machine arises from the necessity of improving the way that pilot candidates are assessed. In this project, the planning for the prototype was developed to improve the accuracy of the results in traditional and computer visual tests.

2.2 Concept Development

The machine must be low cost, occupy little space, configurable for the various tests that are run, and comfortable for user. To avoid delays in project development and unnecessary expenses, we selected a concept according to [12], whom recommends following these steps:

1. Clarify the problem: The problem and the needs must be clear in order to find one or more possible solutions that do not deviate from the project objectives. The main problem in this context is the lack of efficiency during ophthalmological testing.

Table 1 Requirement and engineering features

No.	Requirement	Engineering features
1	Easy assembly and disassembly	Bolted joint design
2	Low cost	Material and maintenance costs
3	Take up little space	Dimensions
4	Comfortable	Ergonomic design
5	Configurable for different test	Adaptability

2. Search externally: The research process for the ophthalmological machine consisted in a bibliographic investigation of projects with similar purposes and information provided by the personnel and doctors of the Ecuadorian Armed Forces.
3. Search internally: From all the ideas presented, the most appropriate ones for the necessities of the project were selected.

Requirements Management The requirements of the device determined the characteristics and engineering parameters given in the Table 1. The engineering characteristics were taken into account to determine and make a prototype that satisfies all the requirements of the device.

Product Outline This diagram represents and describes the idea for every element needed in the machine (see Fig. 1 b). Some elements of the schematic are physical concepts or critical components; however, some other elements continue to be described only functionally.

The product outline of this prototype consists of five elements necessary for its operation. The chair provides support. The table supports the structure. The facial support that fixes the position of the face for testing. The card holder that allows you to change the location and size. The user interface controls the test without changing its basic concept or function.

2.3 System-Level Design

It is important to define the components and subsystems in this phase; the first part of the design is the table and the chair that fulfill the support function. The design of the chair must be adapted for the user's comfort as well as for the different tests.

For this reason, the structure must have bases for monitors that are used in other tests. The facial support must be removable, since not all tests require it and its design must be ergonomic as in [15]. The card support must have three axes of movement and must also be removable, so a CNC was used as a reference. For the implemented software, a Raspberry Pi is the main component to control the system.

Approximate Geometry Once the requirements were obtained, a geometric layout was created in order to test the viability of the physical model. The main problems

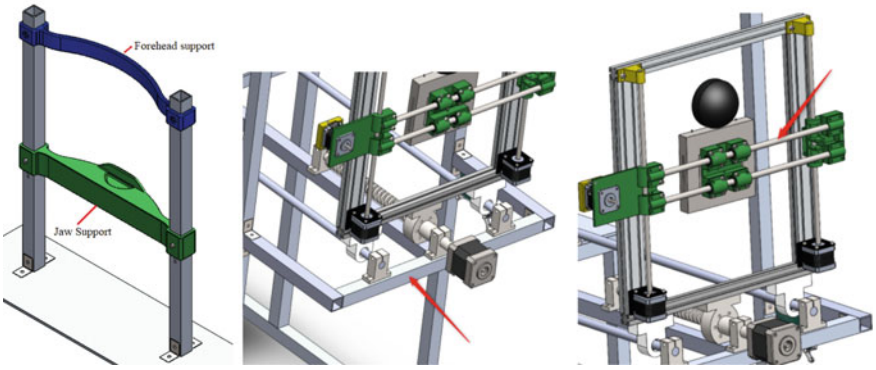


Fig. 2 **a** Face support, **b** motion system, **c** card support

were the position of the facial support and the display for the user. It was possible to resolve applying the final geometric arrangement.

2.4 Detail Design

The main system is decomposed into two subsystems corresponding to different domains: mechanical component and software component. Some components used in the design of this machine, do not need specifications. The chair and table's designs were adapted to standards. The other mechanical elements such as the face support, the movement system, and the card support are described below.

Mechanical component This is the main structure that integrates the different types of visual tests. The structure will locate the necessary elements for the test like monitor, face support, computer, and devices.

- **Face Support** In the Fig. 2a can see the diagram of the support. The dimensions of the elements that compose the facial support were obtained from the facial measurements of those involved in the design, considering a 15% variation. Different measurements can be obtained depending on the user's face.
  The height between chin and forehead will vary due to a sliding link between the tube and the jaw support. It will be graduated every 7 mm with the help of a pin. Both, the support for the forehead and jaw will be design by fused deposition modeling (FDM). It is a deposition of molten material plastic technique by layers. Due to the technique used for manufacturing, polylactic acid(PLA) filament will be used.
- **Motion system—rear support structure** The movement system have three programmable axes were established in, x, y and z, all linear. The X axis allowed the movement of the cards parallel to the facial support to be centered with respect

Table 2 Mesh convergence results

System	Component	Strain error (%)
Facial support	Jaw support	2.06
Motion system	Rear support	1.95
Card support	H.linear guide	1.60

to the patient's view. The Y axis allows the movement of the cards perpendicular to the patient's face, this allows the appropriate distance to be established. The test and Z-move allow to configure the height of the device for testing. This type of metal structures is generally manufactured in commercial structural steel with standard dimensions, in this case, an ASTM A36 steel will be used. See diagram in the Fig. 2b.

– **Card support—Horizontal linear guide** The card support of the manual ophthalmic test cards is located on the x-axis of movement. They can be moved through linear guides (See Fig. 2c). Linear guides are available in standard sizes and are generally made of carbon steel with a chrome surface.

2.5 Testing and Refinement

The testing and refinement phase comprises evaluation before reaching a final prototype. The validation of the mechanical components was carried out using finite elements.

Convergence Analysis

In this study, mesh refinements were carried out in order to study the convergence of the mesh. In this case, an intermediate mesh can be defined and with the default element sizes. The next step is to define an adaptive method, either p-type or h-type. The analysis software determines the rest of the meshing parameters that are best approximate consistent results; due to coarse or fine mesh element transitions that will be generated by the adaptive method. The h-method consists of varying the size of the finite elements according to the local geometry of the component. The p-method is based on the variation of the polynomials used for the analysis. In the properties of the solution method for finite element analysis with the h-type adaptive method, obtained a precision level of 98%, and a number of loops of five were defined in order to find a suitable solution without consuming too many computational resources. The testing of the adaptive method was made for three system with the maximums loads applied facial support, motion system and card support. In Fig. 3 is shown the convergence analysis for the back support.

In order to obtain a better refinement of the mesh and so that the static analysis converges in its results, three critical elements were established. After the analysis, the best was 2.5% as [11] states. The results are given in the Table 2.

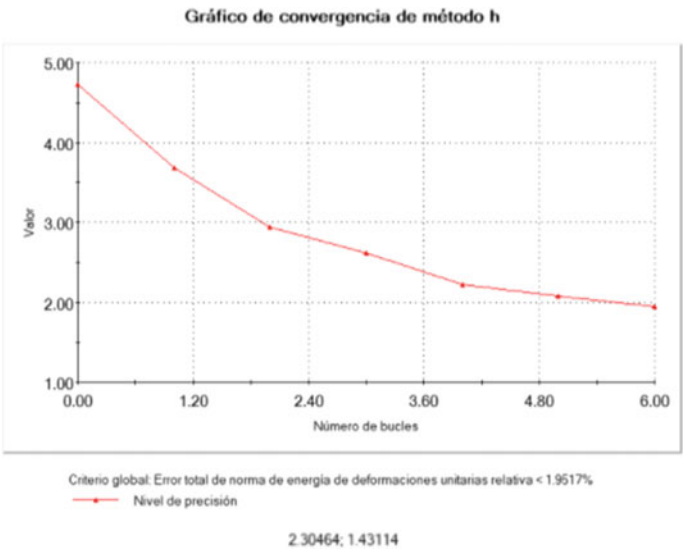


Fig. 3 Convergence analysis in the back support

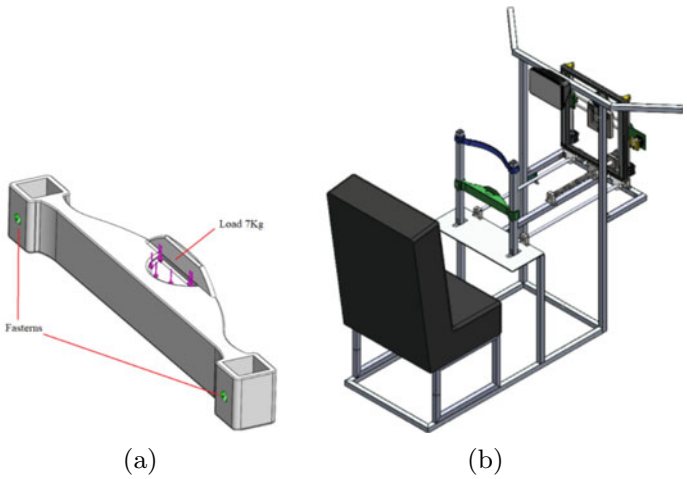


Fig. 4 **a** Load jaw support and **b** final prototype design

Static Analysis To determine the resistance of the facial support where those sections evaluated will support the jaw and forehead. The analysis of the loads and clamps of the jaw support are shown in the following Fig. 4a.

The results of the finite element analysis can be observed, where a maximum stress value is obtained in the area of the fasteners, a value of 1.5 (MPa). With the maximum value present in the element, the safety factor is calculated.

Table 3 Finite element analysis summary by parts

System	Component	Material	T. Strength	FS
Facial support	Jaw support	PLA	47–70 MPa	31.33
Motion system	Rear support	ASTM A36	400 MPa	8.46
Card support	H.linear guide	AISI 1015	385 MPa	198.96

$$FS = 471.5 = 31.33$$

Clearly, the value of the safety factor is greater than 2; therefore, the material assigned for the manufacture of the jaw support is adequate.

The analysis was only carried out on the jaw support since it is the element where the greatest load will be produced, the forehead support only works as a guide. The results of the finite element analysis of all the parts to be analyzed are summarized in the Table 3.

The final product is in the design stage as seen in Fig. 4b, for which the refinement is not yet possible, since there are no physical values to be refined.

3 Conclusions

In this work, we apply the general product development process to design an ophthalmic testing machine. It allows all of the vision tests of the pilot candidates to be evaluated in an integrated way. To satisfy the design points, we established the engineering parameters. Five systems were implemented in the product scheme, chair, table, face support, card holder, and software control. Standard norms were used for the chair and table.

Two support elements were designed with the capability of moving according to patents of ophthalmic equipment and project needs. To determine the strength of the design materials, an adaptive mesh refinement by convergence was performed using the h-method. We designed a prototype of an ophthalmic testing machine that allows traditional and technological tests to be integrated. This machine could improve the accuracy of the results of each tests. It will allow to assessment the participants in an only machine and single place.

4 Future Works

Our project is currently being developed. The goal is to improve the assessment of pilot candidates' visual abilities. The prototype of the ophthalmic testing machine will be integrating traditional and technological tests. We will assess stereopsis and vision colour through experimental studies. Both traditional tests (LANG test and Ishihara test), and two visual tests using virtual reality will be applied.

References

1. Cárdenas-Delgado S, Loachamín Valencia M, Guanoluisa P, Monar-Mejía X (2021) A VR-system to assess stereopsis with visual stimulation: a pilot study of system configuration. Springer, Heidelberg, pp 328–342. https://doi.org/10.1007/978-3-030-68080-0_25
2. Cárdenas-Delgado S, Loachamín Valencia M, Rodríguez Reyes B (2021) VR-test ViKi: VR test with visual and kinesthetic stimulation for assessment color vision deficiencies in adults. Springer, Heidelberg, pp 295–305. https://doi.org/10.1007/978-981-16-4884-7_24
3. Ishihara S (1972) Test for colour-blindness. Kanehara Shuppan Co., Ltd
4. Kim HS, Moon BY, Yu DS, Kim SY, Cho HG (2019) Comparison between stereopsis measured in a natural space and that measured using an apparatus. J Korean Ophthalmic Opt Soc 24(1):71–75
5. Kody F (2012) An integrated aircraft design and performance prediction tool-design, validation, and demonstration. In: 50th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition
6. Lang J (1983) Microtropia. Int Ophthalmol 6(1):33–36
7. Lang J (2019) Lang-stereotest. https://www.lang-stereotest.com/pages/video
8. Momeni Moghaddam H, Kundart J, Ehsani M, Gholami K (2011) The comparison of stereopsis with tno and titmus tests in symptomatic and asymptomatic university students. J Behav Optometry 23(2)
9. O'Connor AR, Tidbury LP (2018) Stereopsis: are we assessing it in enough depth? Clin Experimental Optometry 101(4):485–494
10. Suparyadi D, Yusro M, Yuliatmojo P (2019) Color blindness test by ishihara method based on microcontroller system. KnE Soc Sci :462–470
11. Trivedi A (2007) An adaptive mesh refinement technique for dynamics of solids. University of California, San Diego
12. Ulrich KT (2003) Product design and development. Tata McGraw-Hill Education
13. Wright S, Gooch JM, Hadley S (2013) The role of stereopsis in aviation: literature review
14. Xiong D, Jiang T, Yang L, Wang Y, Guo X, Du J, Bai Y, Su F, Dong W, Lin R et al (2020) Study on pilots' color requirements for see-through displays in desert flight. In: International conference on man-machine-environment system engineering. Springer, Heidelberg, pp 779–787
15. Yoshimura K, Fujii H (2015) Electric optical bench (Japan Patent JP5654322B2)

Knowledge Domain Organization in AEC-AI 4.0 Industry

Multilevel Dynamical Approaches to Knowledge Domain Structures in AEC-AI 4.0 Industry

Carlos Maureira, Héctor Allende-Cid, and José García

Abstract `Knowledge Domain Organization (KDO)` allows us to understand relationships, terms, or topics and the conceptual structure of specific knowledge domains. In complex domains, given by the heterogeneity of actors as well as specific domains/sub-domain features, they also impose further challenges in defining structures/features. In the AEC-AI 4.0 domain, digital and cyber-physical technologies are transforming the sector. However, big unsolved challenges are still foreseen. One of these challenges is the need for development, testing, selection, integration, and orchestration of AI and IoT technology. Here, an integrated approach involving AI, scientometrics, knowledge engineering, and metalearning would enable the development of an integrated operational system for the AEC's life cycle operations. It would provide a dynamic framework for the assessment of digital and cyber-physical technology.

Keywords Architecture engineering and construction · AEC · Artificial intelligence · Internet of Things · Industry 4.0 · Knowledge domain organization · Metalearning · Metaresearch

C. Maureira (✉) · H. Allende-Cid
Escuela de Ingeniería Informática, Pontificia Universidad Católica de Valparaíso, Valparaíso 2374631, Chile
e-mail: carlos.maureira@gmail.com; carlos.maureira@pucv.cl

C. Maureira
Facultad de Ingeniería y Negocios, Universidad de las Américas, Santiago 7500975, Chile

J. García
Escuela de Ingeniería en Construcción, Pontificia Universidad Católica de Valparaíso, Valparaíso 2362804, Chile

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_62

1 Introduction

The architecture, engineering, and construction (AEC) industry is undergoing a process of competitive transformation incorporating artificial intelligence (AI) and developments in Industry 4.0. AEC's life cycle complexity also imposes interesting technological challenges on the development and integration of AI applications into the sector. AEC-AI 4.0 embodies important challenges in scientific terms with important consequences in social, economic, environmental, and sustainability directions. An integrated operational system framework (IOSF) capable of orchestrating a variety of technologies across multiple domains of the AEC's life cycle is a key transformative and enabling component. First, we would briefly introduce main issues in the AEC industry, the building life cycle, and the building information modeling (BIM), which is a central framework for the development of technology . We will offer an overview of the AEC and AI developments, which reveals an important need for integration. After that, we introduce the cybernetics components in the `Knowledge Domain Organization`, developments in the science of science or metaresearch and metalearning as key frameworks along our adaptive methodology [1, 2].

AEC Industry The architecture, engineering, and construction (AEC) industry is one of the oldest in human history and an iconic representation of every civilization. Building information modeling (BIM) is providing a backbone for infrastructure and technologies development [1–4], see Table 1.

Life cycle, BIM, and Technology The AEC is an industry that has lagged behind other industries in terms of productivity gains, stifling national economies' overall growth. Connected sensors, intelligent construction equipment, mobile devices, and software applications could increase productivity, reduce project delays and cost

Table 1 Applications of artificial intelligence in AEC life cycle

Advantages	Planing	Design	Construction	Commissioning and procurement	Operation and management
Visualization, coordinated design, building object properties, specifications, information coordination, information flow	Feasibility study, digital site analysis, pre-design and plan proof, project database creation	3D visualized multidisciplinary design, clash detection and optimization, multidisciplinary coordination, design generation, accurate quantity calculation	Construction site simulation and detailing review, 4D process simulation, schedule control, 5D cost estimation, safety management and quality control	As-built construction inspection, system commissioning	Hetero-system and database integration, Web-based O&M platform establishment 6D facility management

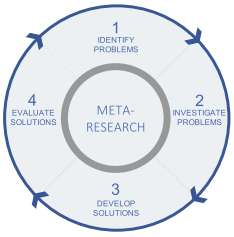
overruns. The AEC's life cycle demands interesting challenges as several interacting domains, operations, technologies, and professionals participate in a well-crucial activity in economic, social, and climate problems. The explosive development of Industry 4.0 is expected to profoundly transform the traditional processes in the AEC. Building information modeling is a game-changing technology that enables end-to-end communication, data exchange, and information sharing among project stakeholders. BIM systems that use Internet of Things (IoT) architectures generate 3D models as part of the BIM process and also as-delivered physical assets [5]. From geographic information to operations and processes throughout the life cycle, including trust and cooperation, can be integrated within the building information model (BIM) framework [6, 7]; but these technologies can also be harnessed to address global climate change and disaster mitigation [8–10]. Therefore, BIM-enabled AEC has many functional features sets although incomplete, to streamline new developments in the industry, especially those involving AI and IoT, and it should be considered as a backbone in the design of any development framework to potentiate the AEC sector. The AEC industry is accelerating its adoption of the most recent technological advances in artificial intelligence (AI) and Internet of Things (IoT) ecosystems. Furthermore, their applications are expected to have a significant impact on the AEC ecosystem and workflow [4, 11, 12]. Similarly, communication and information technologies (CIT), such as edge computing, wireless 5G, and blockchain technologies, will help to boost and reshape the industry. Processes and operations will be significantly transformed in this context, with an increasing emphasis on data generation, analysis, and integration. After that, AEC-AI 4.0 will rekindle interest in a sophisticated workforce with new skills and talents [12, 13]. Edge computing, smart construction objects, cryptography, and data-block technologies, combined with new cultural workforce, social, and sustainable responsibility, will be some of the challenges faced in a competitive transformation process [14, 15]. To characterize and analyze the AEC-AI 4.0 industry workflow, as well as reveal the different computational algorithms used and how the AEC industry can be powered in a fully `Integrated Operational System Framework` (`IOSF`). Here, we expand on research on an efficient global adaptive architectural methodology (M^3) [1, 2]. The proposed development toward an `IOSF` includes the incorporation of *Meta-methodologies* as represented by the metaresearch and metalearning research fields.

Knowledge Engineering. Knowing the conceptual structure of a research field or area is a key aspect in the development, revision, and improvement of the research subjects as well as in the practical advances or applications derived from the research in that particular area. In this context, emerging and established technologies allow us to foster a varied knowledge domains (KD) and further develop solutions to complex problems using technology such as: AI, the Internet of Things (IoT), blockchain, and social networks (SNs). The process of knowledge extraction allows us to better define a field of research as well as monitor and dynamically assess trends, changes, and difficulties. Therefore, there is an increasing interest in how to extract and represent knowledge from research fields [16]. Natural language processing (NLP) becoming broadly recognized as a useful set of methods and tools for the analysis, processing, and mapping of text. A topic, on the other hand, is a central concept in topic modeling,

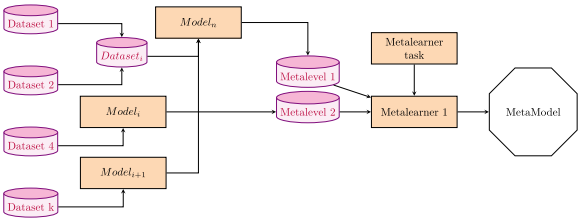
which is typically interpreted as a cluster of words describing a specific semantic meaning. NLP tasks that are more difficult to apply than traditional topic models include text generation, summarization, and translation [17].

2 Metamethods

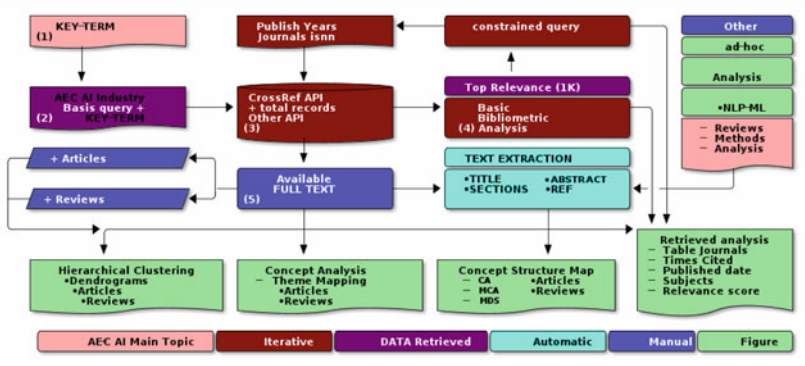
Metaresearch. *Scientometrics and Bibliometrics*. are research methods to analyze scientific literature (see [19–21]); in some ways, these approaches are also regarded as a science about science. Scientometrics allows to assess and analyze the process of knowledge generation [22–25]. In this direction, specific analysis methods such as citation, network, co-words, and content analysis and text mining are used to accomplish their objectives. Studies in this area concentrate on assessing authorship, leading journals, and research organization, but it may also include a content analysis, in some cases derived from words in titles, abstracts, full text in books/journal articles/conference proceedings. Contents analysis also can be evaluated by keywords provided by editors and/or librarians (Fig. 1).



(a) MetaResearch



(b) MetaLearning



(c) Adaptive Methodology

Fig. 1 Metaresearch translational framework [18], metalearning (ML) flow and the adaptive methodology framework for the AEC-AI 4.0 industry, [1]

Table 2 Implementations of ML in AI (modified from [26])

Field	Challenge	Applications	Advantages
FSL	Datasets size constraints	Face recognition, classification, Target detection video synthesis	Low dependence on datasets size, Improved generalization
Robotics	Operation skills	Imitation learning, Cross-domain learning, Quickly adapting	Improve and efficiency of autonomous learning
Unsupervised learning (UL)	Performance	Distribution of unsupervised problems, Noise training	Simplifying UL conversion to supervised learning
Medicine	Progress in the medical field	Medical image processing, drug discovery, diagnosis, question answering	Predicting molecular behavior, Assessment of support samples

Metalearning (ML) seeks to go beyond simple learning by capturing implicit rules in the learning process, tries to gain metaknowledge allowing to autonomously assess, and select the appropriate model, and the parameters can be independently modified to attain additional optimization [26]. Few-shot learning, robot learning, unsupervised learning, and intelligent medicine are some examples of ML applications [26] (see Table 2). Technically, ML not only solves traditional artificial intelligence's data, calculation, and generalization problems, but it also solves machine learning's prediction accuracy and efficiency problems in big data. ML in a general way can be understood as *learning from the learner's generated information*, i.e., learning the metaknowledge of the learned information. Common deep learning models are attempted to implement/learn a model to make predictions, whereas ML is deal with the learning process itself; or in another way, "how to learn a mathematical model faster and better" [26–29].

Although many questions are possible in the context and research approach, we would like to motivate the research with a particular need for defining the `KDO` structure relative to the use of algorithms in the AEC-AI 4.0's life cycle. For example, for the AI/ML algorithms, it would be possible to evaluate: How are they being used? When have they been used? Which requirements and constraints do different algorithms impose in the AEC's life cycle? In particular, we have advanced some research to publication in this context, presenting an initial single macroscopic-level approach to the development of the architectural definition M^3 levels [1, 2].

`Knowledge Domain Organization`, in the AEC-AI 4.0 domain, may allow us to progressively assess the use of computational `Optimization` algorithms, unmet demands, redundancies, deficiencies, and complementarities of processes and operations to orchestrate the development of IOSF. `KDO` structures dis-

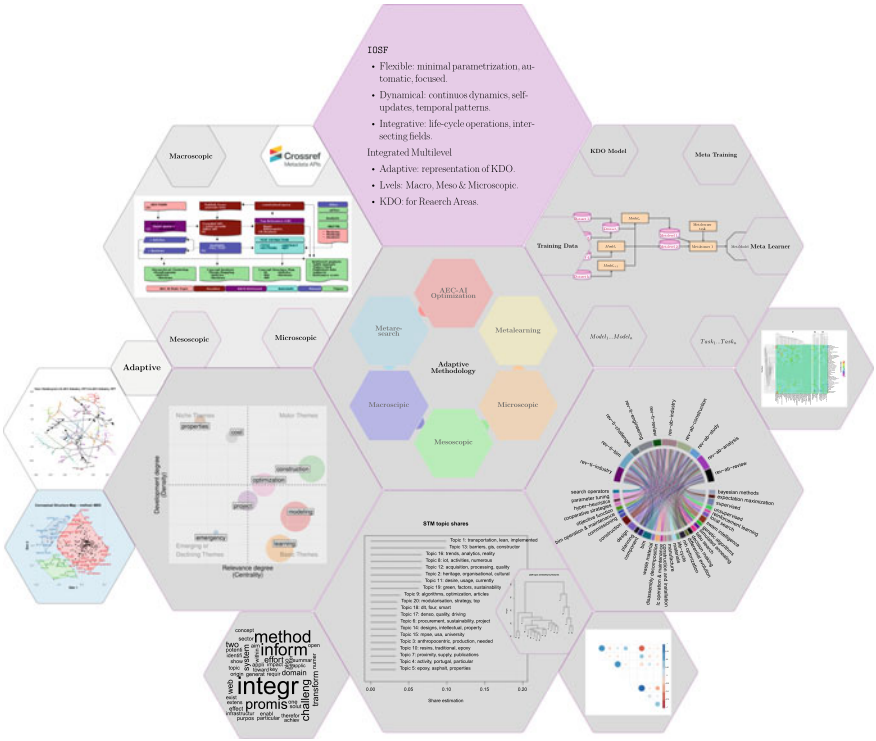


Fig. 2 M^5 Adaptive methodology framework, by extending the M^3 adaptive methodology with metamethods, may allow us to address and integrated operation system framework for the AEC-AI 4.0 industry, [1]

tilled from dynamically formed research corpuses can be defined in conjunction with metalearning and metaresearch to refine an IOSF-(M)-levels initial architecture for the AEC-AI 4.0 domain. The AEC-AI 4.0 IOSF M^5-levels architectural definition will progressively extract `KDO` structures for `Optimization` algorithms in the AEC's life cycle operation. `KDO` structures obtained will be modeled by NLP methods as well as other AI methods, including Bayesian machine learning methods and approaches. Next, we will highlight some requirements and features in the architectural definitions of the M^5-levels initial architecture. We would like to note that either as a dynamically oriented multilevel process, it is highly expected to have sequential formulation of features and requirements in a data-driven process. For heuristic purposes, we will follow a sequence top-to-bottom, but interlink and inverse pathways will be components at each level.

Adaptive methodology. We propose a basic architectural definition to allow three main features of design (Fig. 2): *Flexible*, we propose to implement an adaptive methodology capable of addressing, with minimal parametrization and automatically, the capacity to focus on indented research topics. *Dynamical*, we propose to

incorporate into the architectural definitions the capacity to assess dynamical changes in the organizational knowledge structure to extract temporal patterns. *Integrative*, we propose that the ability to integrate life cycle operations in conjunction with several major intersecting fields of development serve as the basic architectural definition of the proposed `IOSF`. We distill the above architectural requirements into three main stages in defining the KDO: *Macroscopic*, when, first approaching to define the KDO structure, we would like to obtain an overview of the components' topics and the general interconnection between concepts and topics. This macroscopic approach will provide a general assessment of the field of interest. *Mesoscopic*, after obtaining a macroscopic assessment, we would like to obtain more in depth tektological information. Implementing and mesoscopic refinement of the AEC-AI 4.0's KDO structure, we will implement modeling of topics and concepts by machine learning methodologies. In particular, NLP will be used to model, with different perspectives, concepts, and topics. *Microscopic*, a final stage, will assess the microscopic structure in a more oriented predefined approach to obtain tools, methods, agents, and technical specifications in/of the AE-AI 4.0's KDO, constrained to small fractions defined by the relational structure defined in previous stages.

3 Discussion

Previously, we have shown that an adaptive multilevel approach (M^3 methods) can recover the overview of the conceptual mapping in the AEC-AI 4.0 industry when is constrained with bibliometric and scientometric methods [1, 2]. In particular, the key term `Optimization` constrained the research corpus, and the overall conceptual structure recovered is similar to that found by independent methods [4]. In an accompanying paper, we have shown that meso and microscopic levels of analysis can also be derived in the context of the proposed adaptive methodology [2]. Here, we further propose the development of this methodology to incorporate metamethods (Metaresearch and metalearning) to further improve and streamline the adaptive methodology. The proposed methods may allow the formation of an integrated operative system framework for knowledge domain organization in the AEC-AI 4.0 industry. Our proposed $M5$ method (Metaresearch–Metalearning & Macro-Meso-Microscopic) has the potential to generalize across a wide range of complex research domains, allowing adaptive methodology to inform the respective KDO.

Acknowledgements Proyecto 039.414/2021, "Sustainable Business Model Innovation for Construction 4.0".

References

1. Maureira C, Pinto H, Yepes V, Garcia J (2021) Towards an AEC-AI industry optimization algorithmic knowledge mapping: an adaptive methodology for macroscopic conceptual analysis 9:110 842–110 879
2. Maureira C, Allende-Cid H, Garcia J (2022) Optimization in AEC-AI 4.0 industry: multi-level adaptive methodology for knowledge mapping. Submitted
3. Chen C, Tang L (2019) Development of BIM-based innovative workflow for architecture. Eng Construct Proj China April:119–126
4. Darko A, Chan AP, Adabre MA, Edwards DJ, Hosseini MR, Ameyaw EE (2020) Artificial intelligence in the AEC industry: scientometric analysis and visualization of research activities 112(2019):103081
5. Siountri K, Skondras E, Vergados DD (2020) Developing smart buildings using blockchain, internet of things, and building information modeling 12(3):1–15
6. Azhar S (2011) Building information modeling (BIM): trends, benefits, risks, and challenges for the AEC industry 11(3):241–252
7. Di Giuda GM, Pattini G, Seghezzi E, Schievano M, Paleari F (2020) The construction contract execution through the integration of blockchain technology. In: Research for development, pp 27–36
8. Hollberg A, Genova G, Habert G (2020) Evaluation of BIM-based LCA results for building design 109(2019):102972
9. Luo H, Liu J, Li C, Chen K, Zhang M (2020) Ultra-rapid delivery of specialty field hospitals to combat COVID-19: lessons learned from the Leishenshan Hospital project in Wuhan, vol 119, p 103345
10. Alizadehsalehi S, Hadavi A, Huang JC (2020) From BIM to extended reality in AEC industry 116(2019):103254
11. Woo J, Shin S, Asutosh AT, Li J, Kibert CJ (2021) An overview of state-of-the-art technologies for data-driven construction. Lect Notes Civ Eng 98:1323–1334
12. Ceccon L, Villa D (2020) AI-BIM interdisciplinary spill-overs: prospected interplay of AI and BIM development paradigms. In: Springer Tracts in Civil Engineering, Springer, pp 195–217
13. Kipper LM, Iepsen S, Dal Forno AJ, Frozza R, Furstenau L, Agnes J, Cossul D (2021) Scientific mapping to identify competencies required by industry 4.0, vol 64, no January
14. Chong H-Y, Lee C-Y, Wang X (2017) A mixed review of the adoption of building information Modelling (BIM) for sustainability 142:4114–4126
15. Li X, Shen GQ, Wu P, Yue T (2019) Integrating building information modeling and prefabrication housing production 100(December 2018):46–60
16. Leng S, Hu Z-Z, Luo Z, Zhang J-P, Lin J-R (2019) Automatic MEP knowledge acquisition based on documents and natural language processing, pp 800–809
17. Zhao H, Phung D, Huynh V, Jin Y, Du L, Buntine W (2021) Topic modelling meets deep neural networks: a survey. In: Proceedings of the thirtieth international joint conference on artificial intelligence. International joint conferences on artificial intelligence organization, pp 4713–4720
18. Hardwicke TE, Serghiou S, Janiaud P, Danchev V, Crüwell S, Goodman SN, Ioannidis JP (2020) Calibrating the scientific ecosystem through meta-research. Ann Rev Stat Its Appl 7(1):11–37
19. Shi F, Foster JG, Evans JA (2015) Weaving the fabric of science: dynamic network models of science's unfolding structure 43:73–85
20. Zeng A, Shen Z, Zhou J, Wu J, Fan Y, Wang Y, Stanley HE (2017) The science of science: from the perspective of complex systems vol 714-715, pp 1–73
21. Fortunato S, Bergstrom CT, Börner K, Evans JA, Helbing D, Milojević S, Petersen AM, Radicchi F, Sinatra R, Uzzi B, Vespignani A, Waltman L, Wang D, Barabási AL (2018) Science of science 359:6379
22. Larivière V, Haustein S, Börner K (2015) Long-distance interdisciplinarity leads to higher scientific impact 10(3):1–15

23. Sheikhshoaei F (2020) Mapping global knowledge domain, research in information retrieval in medical sciences: a scientometric and evaluative study, pp 1–17
24. Tang M, Liao H, Yepes V, Laurinavičius A, Tupėnaitė L (2021) Quantifying and mapping the evolution of a leader journal in the field of civil engineering 27(2):100–116
25. Zhong B, Wu H, Li H, Sepasgozar S, Luo H, He L (2019) A scientometric analysis and critical review of construction related ontology research 101:17–31
26. Ma P, Zhang Z, Wang J, Zhang W, Liu J, Lu Q, Wang Z (2021) Review on the application of metalearning in artificial intelligence, vol 2021
27. Finn C, Abbeel P, Levine S (2017) Model-agnostic meta-learning for fast adaptation of deep networks 3:1856–1868
28. Xu Z, Cao L, Chen X (2019) Meta-learning via weighted gradient update, 7:110 846–110 855
29. Xu C, Shen J, Du X (2020) A method of few-shot network intrusion detection based on meta-learning framework 15(c):3540–3552

Portable Electronic Dispenser for Personal Hygiene and prevention of COVID-19

Juan Arriola-Condori, Enrique Orihuela-Espinoza, and Michael Cabanillas-Carbonell

Abstract This research proposes the design of a portable electronic dispenser for personal hygiene, which will help reduce the spread of COVID-19, or other diseases or infections that can appear due to lack of hygiene. This dispenser will contain alcohol, will be attachable to the desired arm, either left or right, and will also have a drying mechanism. The disinfectant dispenser and air dispenser will be voice controlled by means of a microphone; the words "alcohol" or "air" must be spoken. When the word "alcohol" is said, the dispenser will begin to expel this liquid toward the direction that is being pointed, be it our hands, body, surroundings, etc. When the word "air" is said, the drying mechanism will be activated, and the air that will start to be expelled will go in the direction that is being pointed, be it to our hands, body, environment, etc. Fourier transform algorithms were used to perform the signal processing. In this work, the MATLAB software was used for the voice recognition of this portable electronic dispenser for the toilet.

Keywords Cleaning · Water dispenser · Voice recognition · Cleaning device · Sensor

1 Introduction

Faced with this problem of the pandemic of COVID-19, this portable electronic dispenser for personal hygiene with voice recognition has been given as a solution, because when people touch a contaminated object, they are at risk of catching COVID-19. This research will allow us to clean ourselves at the time we want without the need to touch a bottle of alcohol, as it will have a built-in alcohol dispenser, which

J. Arriola-Condori · E. Orihuela-Espinoza · M. Cabanillas-Carbonell (✉)
Universidad Privada del Norte, Lima, Perú
e-mail: mcabanillas@ieee.org

J. Arriola-Condori
e-mail: n00048693@upn.pe

E. Orihuela-Espinoza
e-mail: n00118021@upn.pe

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_63

will be attached to the arm, either left or right, and so you can take it to the place you want, without the need to carry something in your hands, pocket, backpack, etc. People infected by COVID-19 present respiratory problems and a series of other symptoms due to the variants that have been developing lately, and it also depends on the organism of the people so that they can get out of this situation, because in some people the virus is defenseless, in others it is more serious and even in others it can be even fatal [1].

To avoid touching things with the hand, just by giving the command, this dispenser will work, thus preventing further contagion of COVID-19. Therefore, this system helps to improve prevention and care for the control of the virus.[2].

For voice recognition systems, one of the greatest difficulties is when there are alterations in its sound, since it makes its recognition difficult. One of the ways in which the voice can be modified is voluntary and involuntary. When it is voluntary, the person, in a lucid manner, alters his or her voice on purpose, such as covering his or her mouth, speaking softly, etc. On the other hand, the involuntary ones occur when the person cannot control such changes, and this can happen when he/she has a cold, a bad throat, hoarseness, etc. There may be other reasons for these involuntary changes, but the ones mentioned above are usually the most common. Another way to modify the tone of voice is artificially, and also by means of electronic equipment that produce these alterations, one of them as an example can be a processor [3].

Because of the mentioned problems that a voice recognition system has, it was for this reason that algorithms were created with the ability to find how much equality there is in the vocalization in order to develop an efficient record, being able to find how many algorithms have better performance, based on its structure and processing to record the patterns [4].

To realize this speech recognition project for a portable electronic dispenser for personal grooming, we have to study the audio signal of the voice and adapt it to the Spanish language; therefore, we will use the automatic speech recognition (ASR) because it will help to manage the user's voice, since the recurrent neural network language models (RNNLM) strongly surpassed the n-gram models that are simpler. Therefore, we will have to sample the speech signal, but we will have to perform two very important tasks, which are training and recognition. It should also be noted that the speech recognition system was not configured in English because there are many differences in their writing and pronunciation, which in comparison with Spanish, these are the same, and help in the recognition of words. And this helps to justify the reason for using this system [5, 6].

2 Methodology

Since the project to be carried out is a portable electronic dispenser for personal hygiene that will be manipulated through voice recognition, the first thing that must be done to perform this communication is to make the signal processing, in this case, the speech of the human being is spoken through a SMD microphone and stored in a

database so that the word that has been said can be recognized, so we must make an interface in communication. The time–frequency transform makes it possible to find out the dynamics of the spectral range over time. The voice signal will be processed, visualized, and plotted in MATLAB, we will make this program communicate with a microcontroller, and for this project, we will use an AVR, the ATMEGA328P of an Arduino Nano, since this microcontroller can receive the information and adapts to this voice recognition system so that we can develop and operate through it a portable electronic dispenser for personal hygiene [7].

When registering and verifying that the word said by the user is correct, in the case of this project will be the words "alcohol" and "air", in order to activate the alcohol jet or the cooler, respectively. In the case of the first sentence, it will activate, with a duration of 2 s, the 5 V mini water pump that will be located in the nozzle of a plastic container filled with alcohol and will expel it through a thin hose. And in the case of the second sentence, the cooler for drying hands or the place where the device is voluntarily pointed by the user will be activated for a duration of 4 s. The Fast Fourier transform (FFT) helps the registration of sound signals and to solve a large number of variables that hinder speech recognition, and through artificial intelligence is that we can also get to do speech recognition [8].

When the training has been done, the voice recognition system has been strengthened, and the words "alcohol" and "air" have been stored in all possible environments, the electronic circuit of the electronic dispenser will be made, which in this case will be a timer, because at the moment of pronouncing the respective word to activate the portable electronic dispenser for personal hygiene, the microcontroller will give the order to the circuit, and by means of a clock, it will be able to carry out only 1 of the 2 following actions, activate the 5 V mini water pump for a time of 2 s, or activate the cooler for a time of 4 s. The ASR as it includes a dataset, we make decoding parameters for a utility evaluation, and the retraining is superior compared to training with random weights. We embraced a RNNLM as a conventional semi-supervised method for comparison [9].

MATLAB software was used, since this program allows to perform the work of voice recognition for the portable electronic dispenser for personal grooming in front of COVID-19, also an AVR ATMEGA328P microcontroller of an Arduino will be used, in order to carry out the electronic part of this portable electronic dispenser for personal grooming in front of COVID-19, and also the Fourier formulas will be used, since these will help us to perform the concepts for signal processing. From another point of view, the Wavelet Transforms (WT) are also a way to study the signal in time and frequency, it is a simple way to show a complex signal, since it decomposes it into small waves of lower resolution and different frequencies, managing to find them in the original signal. On the other hand, the Fourier transform discovers the frequency of the disturbance but without being able to find the time in which it is found [10].

We have had to create the MATLAB and C++ codes for the AVR ATMEGA328P microcontroller to carry out this project, and the materials have had to be put together in such a way that they fit the left or right arm. The decoding procedures help to obtain the cost-effectiveness closer to the new generation auto-regressive models, but the

non-auto-regressive ones only perform the main decoder operations with a constant number "*k*" times [11].

The procedure that we will do at this point is about what steps we had to do to perform speech recognition in this portable electronic dispenser, and for this, we will do some speech recognition tests in MATLAB, and then a flowchart where we will observe how we have programmed the speech recognition. The decoding makes an acoustic sequence through the network, continuous selection of the word with the highest score at each moment of time [12, 13] (Figs. 1, 2 and 3).

Next, in reference to the programming code flowchart, a schematic was carried out as a base structure in order to understand the parts of the project in design condition (Fig. 4).

Therefore, it was taken into consideration which components can be used for the portable electronic dispenser, being designed in the Fritzing program with a determined order [14], the components to appreciate are (Fig. 5):

- Arduino ProMini.
- Battery Charger Module.
- Voltage Booster Module.
- Microphone Module.
- Lithium Battery.

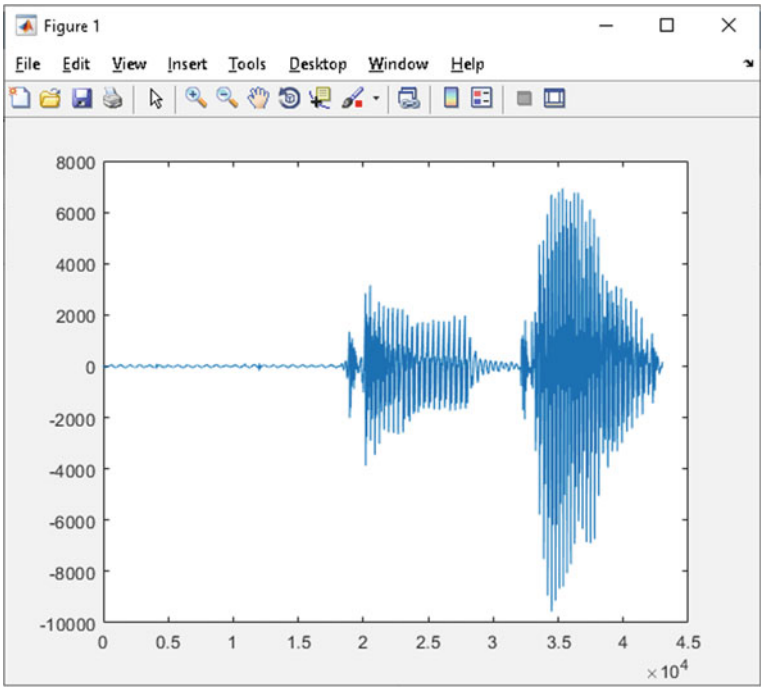


Fig. 1 Voice signal when saying the word: alcohol

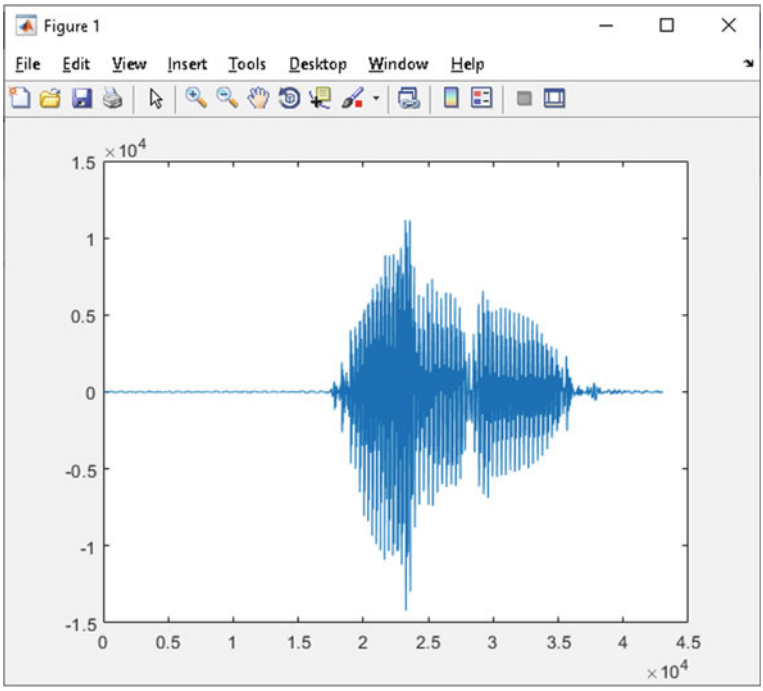


Fig. 2 Voice signal at the moment of saying the word: air

– Switch.
– DC Motor.
– 5v Fan.

Consequently, the 3D model was designed in the FreeCAD program with the order of each component in order to have a more concrete idea about the portable electronic dispenser [15]. With this project, the intention is to strengthen individual and public cleaning habits, since it is an alternative of easy use and application with greater resistance and portability for people (Fig. 6).

3 Results

When performing the simulations in MATLAB for speech recognition, it can be seen that there is a 1–3 s delay for the code to be compiled. When performing the speech recognition, first training had to be performed to identify the words that will operate the portable electronic dispenser for personal hygiene, which are: alcohol and air. These trainings first had to be loaded into a "mat-file", as this will be our database.

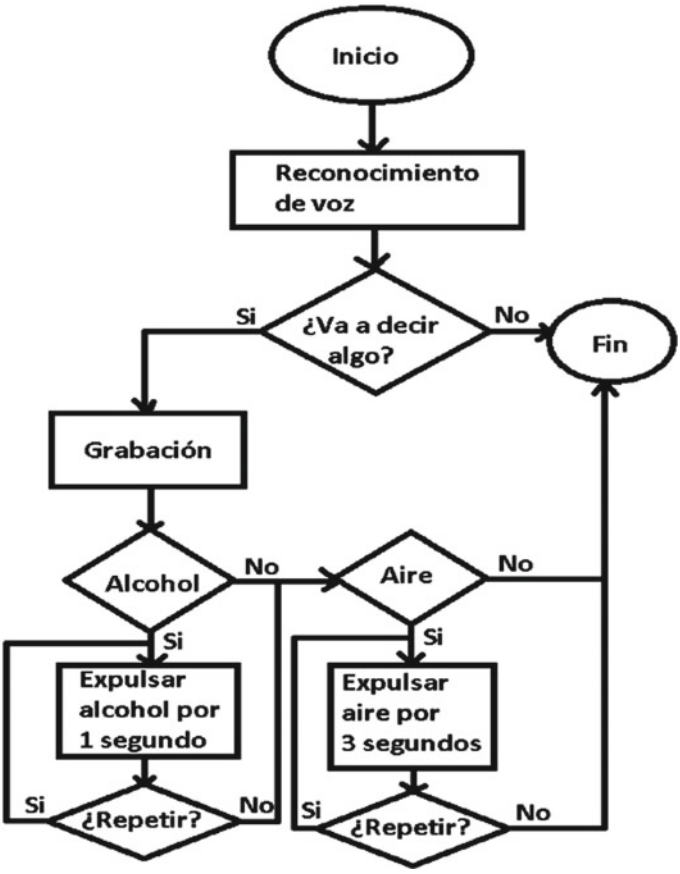


Fig. 3 Speech recognition flowchart

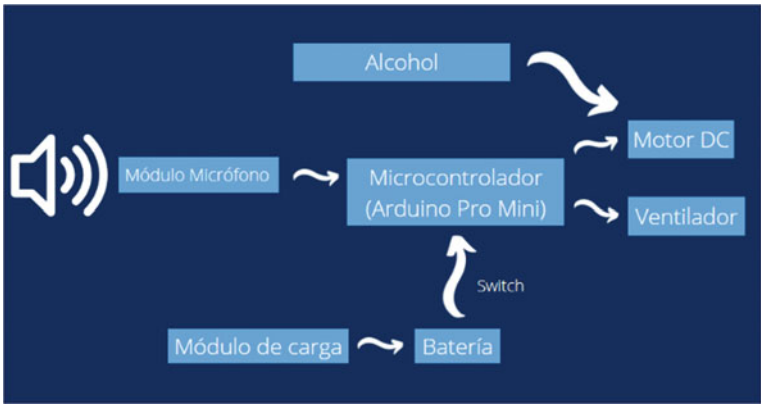


Fig. 4 Block diagram

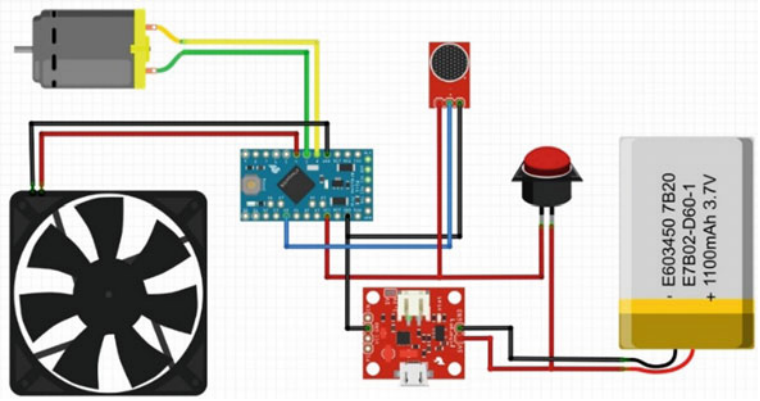
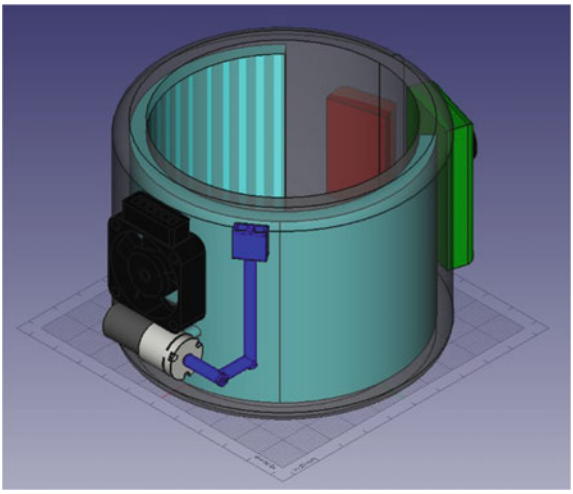


Fig. 5 Electronic design

Fig. 6 3D model design



The following are the plots of the Fast Fourier transform that were performed on the voice signals (Figs. 7 and 8).

And by doing so, the real values were found to the Fast Fourier transform of these signals, in order to obtain their linear indices by means of a MATLAB function called "find". These linear indexes, at the time of training, the following values were obtained according to the recorded word (Tables 1 and 2).

As can be seen, the values of the indexes of each word are different; these results were obtained at the time of training. These differences exist because they do not have the same tone, frequency, and signal; since, they are different words and also have different sounds. To recognize which word it is, we had to calculate the difference between the values of the indexes stored in the database, and those that will be

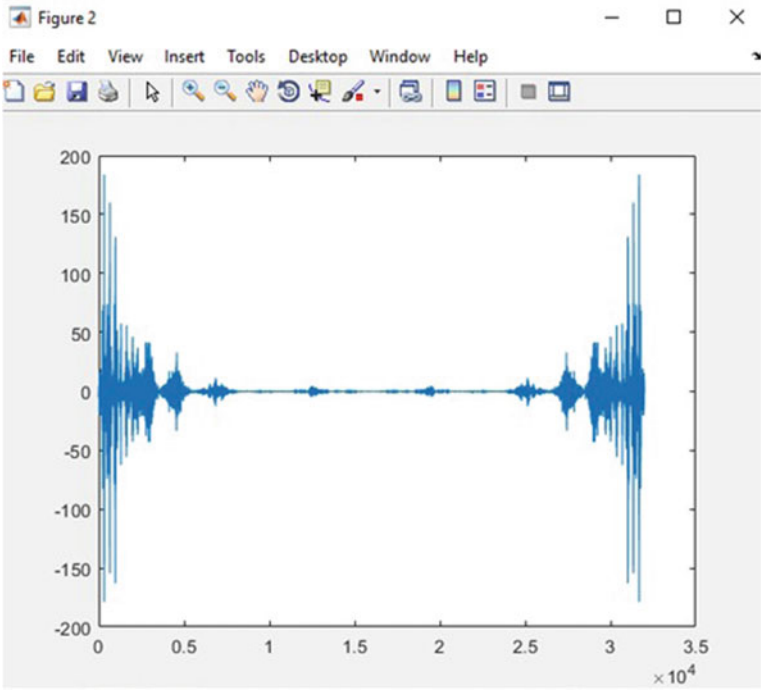


Fig. 7 Fourier transform to the word signal: alcohol

recorded in real time, and the minimum value of these two is the one that will help us determine which word is being said (Figs. 9 and 10).

Figures 11 and 12 show the moment when the words "alcohol" and "air" are recorded in real time, so that this signal can then be compared with the data from the recordings already stored in our database.

More recordings were then made to save them in the database, in order to strengthen the portable electronic dispenser project for personal hygiene. And when this was done, it was possible to have the optimal recognizer that was desired at the beginning of the work. This is the meaning of each letter and word shown in the Datos Workspace (Figs. 11 and 12): "C" is the number of order and trainings that were made with the words to be recognized. The word "alcohol" is number 1, and "air" is number 2. "F" are the values of the linear indexes of the words "alcohol" and "air" that are stored in the database. The first five numbers are for the word "alcohol", and the remaining numbers are for the word "air". "*f*" is the value of the linear index that was obtained at the time of the real-time recording. "*D*" is the subtraction of the linear index values that have "*F*" and "*f*". "*d*" is the last linear index value obtained in the subtraction between "*F*" and "*f*". "*i*" is the total number of index values in the database. ind: is the order number of the minimum value obtained in "*D*". Record_number: represents the order number that was identified when doing

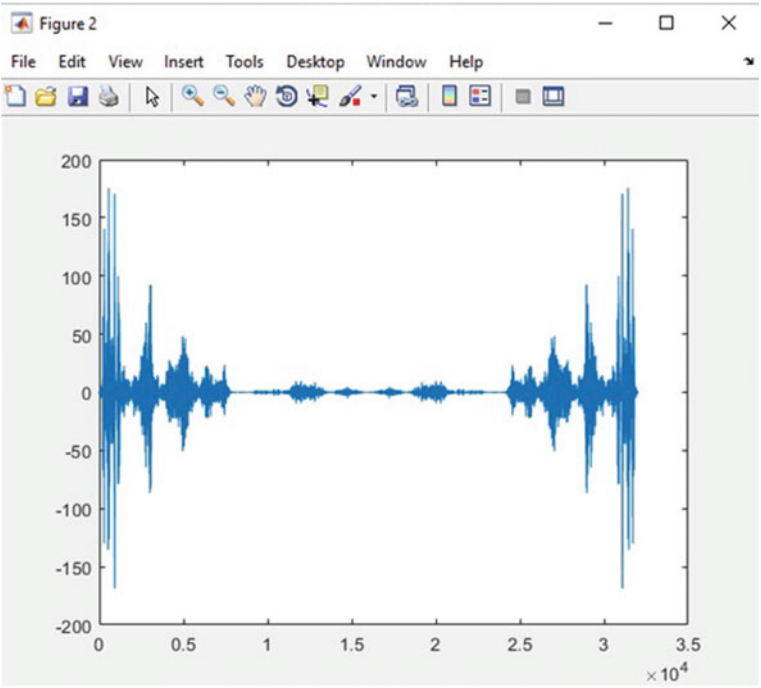


Fig. 8 Fourier transform to the signal of the word: air

Table 1 Index values of the word: alcohol

Word indexes: alcohol	Index value
Índice 1	2961
Índice 2	947
Índice 3	1026
Índice 4	990
Índice 5	3213

Table 2 Index values of the word: air

Word indexes: alcohol	Index value
Índice 1	575
Índice 2	922
Índice 3	565
Índice 4	549
Índice 5	1141

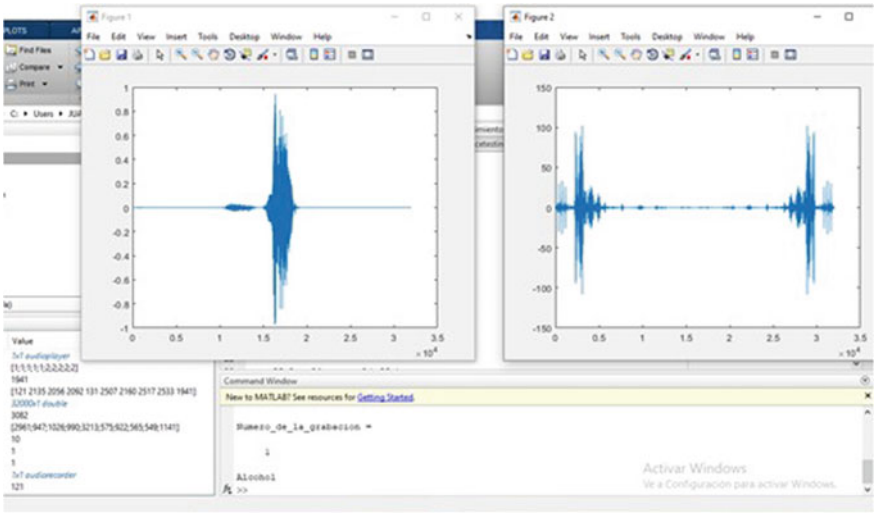


Fig. 9 Real-time recognition of the word “alcohol”

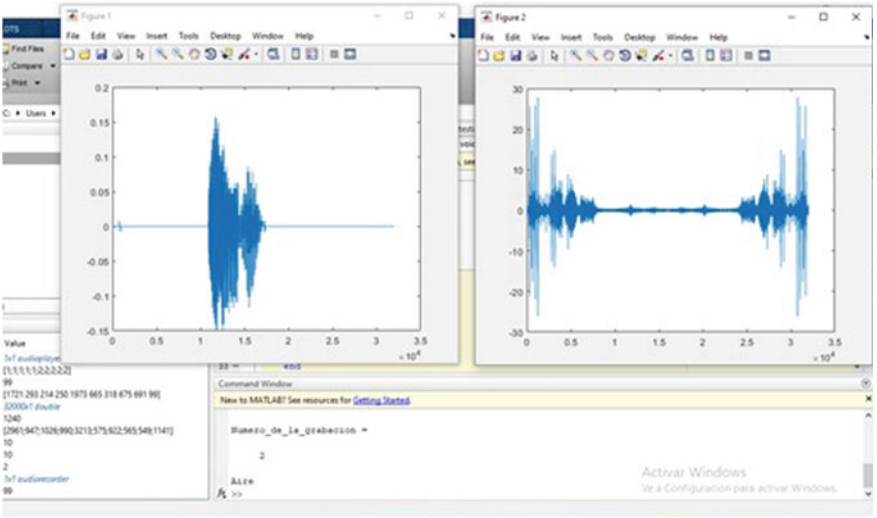


Fig. 10 Real-time recognition of the word “air”

the voice recognition. sm: is the minimum value that was obtained when doing the subtraction between “*F*” and “*f*”.

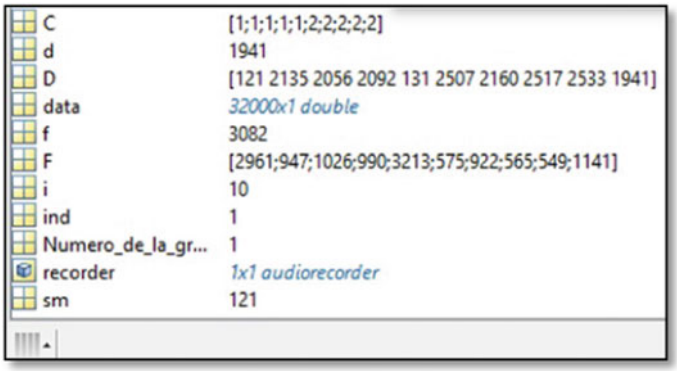


Fig. 11 Workspace data of the recordings stored in the database, together with the real-time recording (f) of the word: alcohol

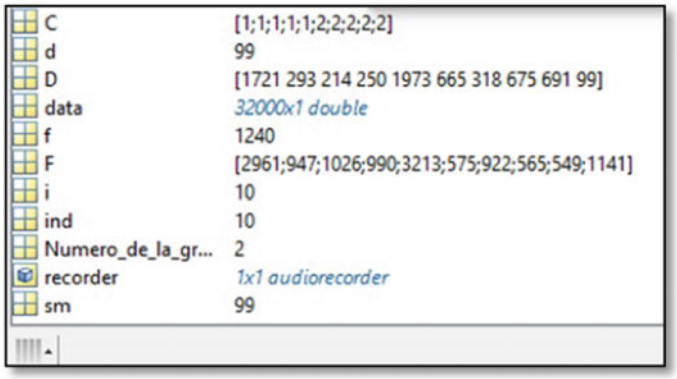


Fig. 12 Workspace data of the recordings stored in the database, together with the real-time recording (f) of the word: air

4 Conclusions

The project model developed was very significant to provide a solution to COVID-19 infections, since it ensures a higher level of cleanliness.

The general objective is to develop an effective voice recognition procedure, in any type of environment and circumstance, for a portable electronic dispenser for personal hygiene. A signal processing that achieves a real-time voice recognition with a duration of 2 s was achieved. With respect to the first specific objective, an optimal system was achieved, since at the time of training, we were able to make it more efficient at the time of voice recognition. The trainings that were performed also had a recording time of 2 s, since it is desired to recognize the voice in the shortest time possible to avoid other sound interferences. With respect to the second

specific objective, the noise was eliminated but not completely, since at the moment of doing the voice recognition in real time sometimes external voices were filtered if the noise was very loud, but with the low noises there were not so many problems. And this noise elimination was achieved thanks to the Fourier transforms. With respect to the third specific objective, it was concluded that the AVR ATMEGA328P microcontroller helps the electronic circuit to be reduced in a certain way, in addition, for there to be a response from these components, there will always be a latency for them to respond to their objectives.

References

1. World Health Organization, "Coronavirus," 2020. [En línea]. Available: https://www.who.int/health-topics/coronavirus#tab=tab_1. [Último acceso: 08 Jan 2022]
2. Gobierno del Perú, "Minsa actualiza protocolos de bioseguridad para la prevención y control de la COVID-19 en el Perú," 4 noviembre 2021. [En línea]. Available: https://www.gob.pe/institucion/minsa/noticias/551627-minsa-actualiza-protocolos-de-bioseguridad-para-la-prevencion-y-control-de-la-covid-19-en-el-peru/. [Último acceso: 08 Jan 2022]
3. Martínez GA, Aguilar G (2013) Reconocimiento de voz basado en MFCC, SBC y Espectrogramas. Ingenius. Revista de Ciencia y Tecnología (10):12–20. https://doi.org/10.17163/ings.n10.2013.02
4. De Luna CA, Martínez JC, Mora M (2006) Reconocimiento de Voz con Redes Neuronales, DTW y Modelos Ocultos de Markov. Conciencia Tecnológica (32):13–17
5. Oropeza J (2006) Algoritmos y Métodos para el Reconocimiento de Voz en Español Mediante Sílabas. Computación y Sistemas 9(3):270–286
6. Li K, Xu H, Wang Y, Povey D, Khudanpur S (2018) Recurrent neural network language model adaptation for conversational speech recognition. In: Proceeding Interspeech, pp 3373–3377. https://doi.org/10.21437/Interspeech.2018-1413
7. Morales M, Echeverry JD, Orozco ÁÁ (2007) Reconocimiento de emociones empleando procesamiento digital de la señal de voz. Scientia Et Technica 8(37):145–150. https://doi.org/10.22517/23447214.4019
8. Esparza ME, Avalos JB (2003) Reconocimiento de voz. Conciencia Tecnológica, nº 23
9. Karita S, Watanabe S, Iwata T, Ogawa A, Delcroix M (2018) Semi-supervised end-to-end speech recognition. In: Interspeech, pp 2–6. https://doi.org/10.21437/Interspeech.2018-1746
10. San Juan E, Jamett M, Kaschel H, Sánchez L (2016) Sistema de reconocimiento de voz mediante wavelets, predicción lineal y redes backpropagation, Ingeniare. Revista Chilena de Ingeniería 24(1):8–17. https://doi.org/10.4067/S0718-33052016000100002
11. Chen N, Watanabe S, Villalba J, Żelasko P, Dehak N (2021) Non-autoregressive transformer for speech recognition. IEEE Signal Process Lett 28:121–125. https://doi.org/10.1109/LSP.2020.3044547
12. Audhkhasi K, Ramabhadran B, Saon G, Picheny M, Nahamoo D (2017) Direct acoustics-to-word models for english conversational speech recognition. In: Interspeech, pp 959–963. https://doi.org/10.21437/Interspeech.2017-546
13. Singla SK, Yadav RK (2014) Optical character recognition based speech synthesis system using LabVIEW. J Appl Res Technol 12(5):919–926. https://doi.org/10.1016/S1665-6423(14)70598-X

14. Solis JF, Yáñez C, Suárez S (2011) Reconocimiento automático de voz emotiva con memorias asociativas Alfa-Beta SVM. Polibits 44:19–23. :https://doi.org/10.17562/PB-44-3
15. Karita S, Yalta NE, Watanabe S, Delcroix M, Ogawa A, Nakatani T (2019) Improving transformer-based end-to-end speech recognition with connectionist temporal classification and language model integration. In: Interspeech, pp 1408–1412. https://doi.org/10.21437/Interspeech.2019-1938

ANN Model for Two-Way Shear Capacity of Reinforced Concrete Slabs Without Shear Reinforcements

Nermin M. Salem and A. Deifalla

Abstract This study aims to develop a reliable model for the two-way shear capacity of reinforced concrete slabs with FRP reinforcements, which is a complicated problem with many effective parameters. In this study, an ANN capacity model was developed and proposed. The model predicted the capacity accurately with respect to the available models. In addition, the effect of the main parameters on the capacity using the proposed model was investigated and thus provided an insight on the influence and interrelation of effective parameters on such a problem which can help further design code development.

Keywords Two-way shear · Slabs · GFRP · CFRP · FRP · ANN

1 Introduction

The two-way shear capacity of concrete slabs with FRP reinforcements is a sudden failure [1–3]. Thus, avoiding such catastrophic failure requires advanced methods for assessing the reliability of design models [4]. In the previous decade, several studies have been conducted based on the usage of Artificial Neural Networks (ANN) in solving various real-life construction problems such as structural analysis and design, building materials, and solving management problems. ANN can model complex behavior in a much better way compared to other modeling techniques. This optimization problem lies in the ANN learning process and could be solved with the selection of the appropriate training algorithm [5]. For the two-way shear capacity of FRP-reinforced concrete slabs, a hand full of studies attempted to develop an AI model; in a study by Metawally, an ANN model was developed based on TRAINLM for the mapping between the input parameters and the two-way shear capacity, using

N. M. Salem (✉)
Electrical Engineering Department, Future University in Egypt, Cairo, Egypt
e-mail: nfawzy@fue.edu.eg

A. Deifalla
Structural Engineering and Construction Management Department, Future University in Egypt, Cairo, Egypt
e-mail: ahmed.deifalla@fue.edu.eg

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_64

a dataset composed of 59 records [6]. In a study by Naderpour et al., an ANN model was developed using a back-release network. The model was trained using a dataset composed of 50 records [7]. In a study by Vu and Hoang, two models were developed an ANN model and a support vector machine (SVM) model using a dataset composed of 82 records [8]. It is worth mentioning that the accuracy of an ANN model mainly depends on the used database in terms of size, variety, and quality [9, 10]. Thus, in the current study, a larger database was targeted. In this paper, an ANN model was developed, which was improved by implementing a momentum and an adaptive learning rate.

2 Development of Artificial Neural Networks (ANN) Model

The use of ANN in predicting the capacity of reinforced concrete slabs with FRP reinforcement has been rarely presented, where only hand full of studies are available in the published literature. ANN was originally inspired from the human biological nervous systems; ANN consists of many processing elements called neurons. Neurons are connected to each other by links called weights. Initially, during training, weights are randomly assigned; afterward, they are corrected using the predicted values compared with the real values; these errors are then backpropagated through the network; and the final weights are recalculated to minimize the measured errors. The process of ANN training could be summarized in three main steps: (1) initialization of internal parameters; (2) model evaluation; and (3) updating the internal parameters to find the optimal points. The network weights are updated through backpropagation, and it is a method of fine-tuning the weights of an ANN-based on the error rate obtained in the previous training iteration. The powerfulness of ANN is the ability to capture complex relationships between input and output parameters without any prior information about the nature of these interactions. However, its implementation based on the traditional gradient descent method for the ANN training process could result in large errors and suffer from several local minima [11]. Thus, the proposed ANN model is trained using gradient descent with two main modifications: adding a momentum term and using an adaptive learning rate. The proposed model is trained based on the available experimental data of a total of 189 records [12–14]. AI had been used in various applications such as [16–18].

ANN has four main training methodologies: gradient descent (ANN-GD), quasi-Newton method (ANN-QN), Levenberg–Marquardt (ANN-LM), and conjugate gradient (ANN-CG). The challenge lies in choosing the suitable training methodology as a given training technique may be suitable for a given problem but fails in another scenario [5]. Traditional ANN-GD shows promising reliable results versus the other techniques [11] where it reduces the errors between the real experimental output and the predicted output from the trained model [15] and can easily adapt to the complex behavior of the training data. ANN-GD is considered the most popular optimization technique in ANN, it aims to find a local minimum of a differentiable function, and in other words, it is to find the coefficients that minimize the error

function as much as possible. It measures the change in all weights regarding the change in error [19].

The traditional ANN-GD is a training model that updates its weights and biases in the direction of the negative gradient of the performance function is such that:

$$w^{(i+1)} = w^{(i)} - \nabla f\left(w^{(i)}\right)\eta^{(i)} \tag{1}$$

where $w^{(i)}$ is the weights in the ith iteration, $\nabla f\left(w^{(i)}\right)$ is the loss function gradient f according to $w^{(i)}$, and $\eta^{(i)}$ is the learning rate of the ith iteration. Learning rate η is a hyperparameter in the training process as it controls the changes in the model, i.e., how quickly the model is adapted to the given problem, with respect to the measured error in each training iteration, and it may be the most important hyperparameter while configuring the network, as choosing a too large value for η may lead to a model that converges too quickly to a suboptimal model and also choosing a too small value may lead to an unstable model. The challenge is to find the appropriate η value [19]. In this model, an adaptive learning rate is used, i.e., the default value is 0.01 that is usually used by the traditional ANN-GD. Afterward, it is adjusted during training according to the performance measured. At the end of each training iteration, based on training, if the performance decreases toward our goal, then the learning rate is increased by the factor $\eta_{\text{inc}} = 1.05$. If performance increases by more than a specific factor, which is predefined in the training phase, it is decreased by a factor $\eta_{\text{dec}} = 0.7$. The advantage of using an adaptive learning rate is faster and more stable training. The used training methodology is gradient descent with momentum and adaptive learning rate backpropagation. The momentum term (mc) helps in increasing the speed of convergence, without compromising the model capability in solving nonlinear problems, where it is computed through the derivatives of performance (p) with respect to the weight and bias parameters x for each training iteration i. Each parameter is adjusted according to gradient descent with momentum:

$$d_{x(i)} = \text{mc}^*d_{x(i-1)} + \frac{\eta^*\text{mc}^*\text{dp}}{d_x} \tag{2}$$

where $d_{x(i-1)}$ is the previous change on the weights or bias, mc is the momentum coefficient, x is the weights and bias, and dp is the derivative of the performance.

The used training methodology can train any network if its weight, inputs, and transfer functions have derivative functions. The ANN model learns from the experimental database of reinforced concrete slabs with FRP reinforcements to determine the two-way shear capacity. The model consists of an input layer, one hidden layer with 11 neurons, and an output layer. The model is trained with five inputs: column dimensions C, effective depth d, concrete compressive capacity f'_c, flexure reinforcement ratio ρ, and young's modulus (E). The output is the prediction of the two-way shear capacity of FRP V. The training summary of the ANN algorithm is shown in Table 1.

Table 1 ANN training summary

Parameter	Description
Neurons in input layer	5
Neurons in output layer	1
Hidden, output layer activation function	Sigmoid, Linear
Cost function	*MSE*
Number of hidden layers	1
Number of training iterations	*9*

Our trained model will stop training if any of the following conditions occur; (a) maximum training iterations is reached, (b) maximum time is reached, or finally, (c) performance goal is reached. In our model, the stopping criteria were the third condition. The model was trained using a dataset composed of 189 records. For best performance evaluation, the dataset is divided into two main parts as recommended by [19]: training set containing 70% of the total data, and the remaining 30% was used for validation and testing. The dataset has been used without any post-processing.

3 Comparison between Proposed Model and Selected Models

To evaluate the effectiveness of the proposed models, the following statistical measures were calculated for the proposed model and the selected models where average values of the ratio between the measured and calculated capacity, the root mean square error (RMSE), the coefficient of variation (CoV), and the lower limit based on a one-sided 95% confidence interval. Figure 1 shows the comparison between the selected models and proposed models based on these four statistical measures, and the proposed models are more accurate, consistent, and reasonably safe for design, where the ANN model had an average, coefficient of variation, root mean square error, and lower 95% value of 0.97, 20%, 1.1, and 0.93, respectively. Figure 2 shows a parametric study for the capacity calculated by each model versus the values of the effective parameters. For the effect of the effective depth, the variation of the proposed models' predictions is close to that of the CSA and JSCE up to an effective depth value of 200 mm. After 200 mm value for the depth, the variation of the proposed models' predictions is significantly different. There is a reduction in the capacity with the increase in depth (i.e., size effect). While other models including the CSA, Ju, and JSCE show no size effect. Moreover, Ju model prediction increases with the increase in size depth even after 200 mm value for the effective depth. For the flexure reinforcement ratio, the variation of the proposed models' predictions shows an almost flat relationship. On the other hand, other models including Ju, CSA, and JSCE show a clear nonlinear variation. Such effect could be attributed to the significant reduction in the transversal resistance, thus diminishing the dowel

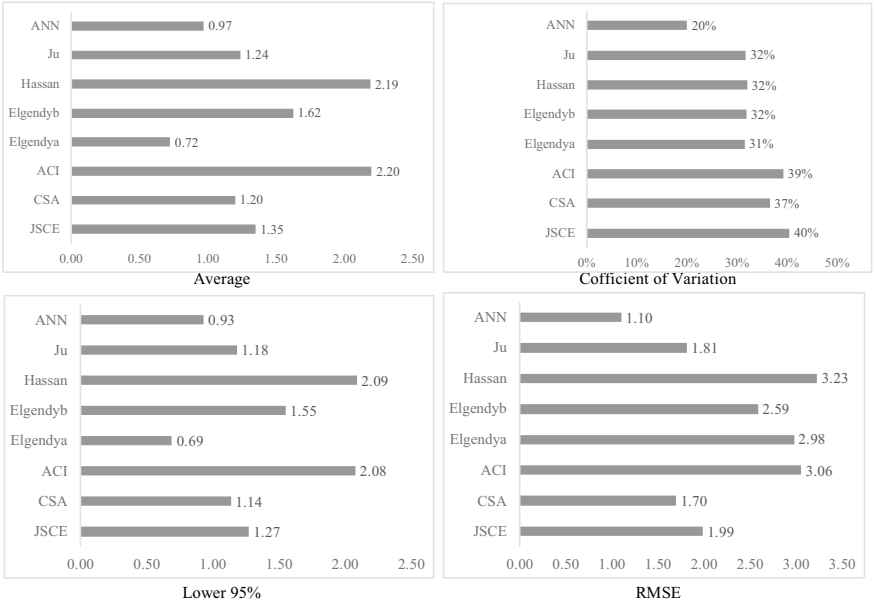


Fig. 1 Validation of the proposed ANN model

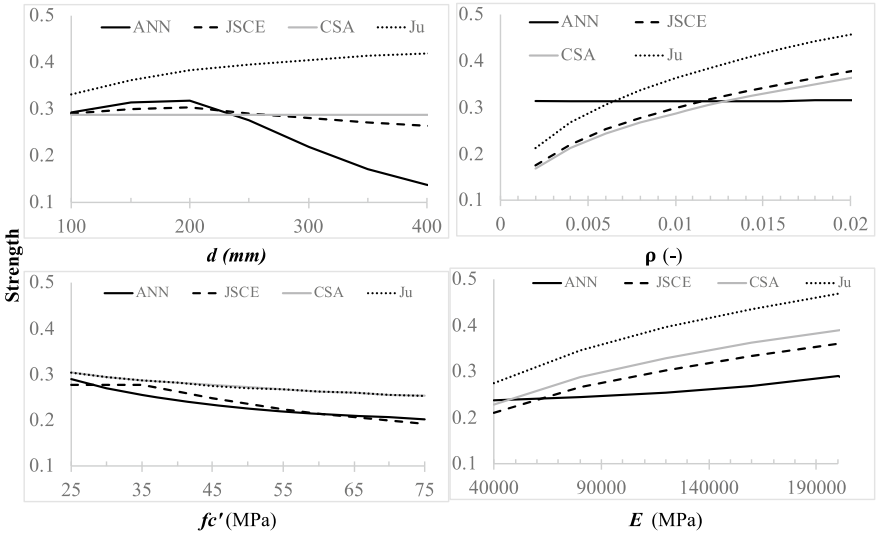


Fig. 2 Parametric study for the capacity calculated using all model versus effective parameters

action component. For the effect of the concrete compressive capacity, the variation of the Ju and CSA model predictions is nonlinear and different from that of the JSCE. Since the JSCE prediction is related to $\sqrt[2]{f_c'}$and the Ju and CSA are related to $\sqrt[3]{f_c'}$, the proposed models show a completely different variation, which is not the $\sqrt[2]{f_c'}$ or the $\sqrt[3]{f_c'}$[3]f_c'variation. For the effect of E on the variation of the capacity, the proposed models show a much lower variation than all other models as well as much lower values. From the proposed models, it can be shown that the effect of size and concrete compressive capacity is dominant. This might be due to the lower dowel action effect and young's modulus values of FRP bars.

4 Conclusions

An ANN model was developed and implemented in a parametric study. The model was found to be more accurate, consistent, and reasonably safe for design compared to select models. Both models were capable of modeling the complex behavior of reinforced concrete slabs with FRP reinforcement under two-way shear as well as include the parameter uncertainty. A parametric study is conducted for the effect of various parameters. Although conclusions are limited to the range of the experimental database, however, the following conclusions were reached at: (1) The flexure reinforcement ratio was found to have a small effect on the capacity, which could be attributed to the significant reduction in the transversal resistance of FRP reinforcement, thus diminishing the dowel action component. (2) The effect of concrete compressive capacity was found to be different than the traditional $\sqrt[2]{f_c'}$ or the $\sqrt[3]{f_c'}$ variation. (3) The effect of the size and concrete compressive capacity was found to be dominate on the capacity, which might be due to lower dowel action effect and young's modulus values of FRP bars.

References

1. Deifalla AF (2021) A mechanical model for concrete slabs subjected to combined punching shear and in-plane tensile forces. Eng Struct, Elsevier 231(Mar)
2. Deifalla AF (2021) A strength and deformation model for prestressed lightweight concrete slabs under two-way shear. Adv Struct Eng 1–12. https://doi.org/10.1177/13694332211020408
3. Deifalla AF (2020) Strength and ductility of lightweight reinforced concrete slabs under punching shear. Structures 27:2329–2345. https://doi.org/10.1016/j.istruc.2020.08.002
4. Deifalla A (2022) Punching shear strength and deformation for FRP-reinforced concrete slabs without shear reinforcements. Case Stud Construction Materials 17(July):2022
5. CSA (2012) Design and construction of building structures with fiber reinforced polymers (CAN/CSA S806–12). Rexdale, ON, Canada
6. ACI-440 (2015) Guide for the design and construction of concrete reinforced with FRP Bars (ACI 440.1R-15). ACI, Farmington Hills, Michigan, USA

7. Hassan M, Fam A, Benmokrane B (2017) A new punching shear design formula for FRP-reinforced interior slab-column connections. In: 7th International conference on advanced composite materials in bridges and structures. Vancouver, British Colum
8. Ju M, Ju JW, Sim J (2021) A new formula of punching shear strength for fiber reinforced polymer (FRP) or steel reinforced two-way concrete slabs. Composite Struct 258(Mar):113471. https://doi.org/10.1016/j.compstruct.2020.113471
9. Sorensen J, Baravalle M, Kohler J (2019) Calibration of existing semi-probabilistic design codes. In: 13th International conference on applications of statistics and probability in civil engineering, ICASP13. South Korea, Seoul National University
10. Naderpour H, Shali MG, Fakharian P (2016) Application of artificial neural network in punching shear capacity of FRP reinforced slabs. In: The 8th national conference on concrete. Tehran, Iran. https://doi.org/10.13140/RG.2.2.26289.79201
11. Vu D-T, Hoang N-D (2016) Punching shear capacity estimation of FRP-reinforced concrete slabs using a hybrid machine learning approach. Struct Infrastruct Eng 12(9):1153–1161
12. Ikumi T, Galeote E, Pujadas P, Fuente Al, López-Carreño RD (2021) Neural network-aided prediction of post-cracking tensile strength of fibre-reinforced concrete. Comput Struct 256:106640
13. Mangalathu S, Shin H, Choi E, Jeon J-S (2021) Explainable machine learning models for punching shear strength estimation of flat slabs without transverse reinforcement. J Build Eng 39:102300
14. Nguyen T-A, Ly H-B, Mai H-VT, Quan TV (2021) On the training algorithms for artificial neural network in predicting the shear strength of deep beams. Complexity 2021(3):1–18. https://doi.org/10.1155/2021/5548988
15. Lv Y, Lê Q-T, Bac BH, Bui X-N, Nguyen H, Nguyen-Thoi T, Dou J, Song X (2020) A comparative study of different machine learning algorithms in predicting the content of Ilmenite in titanium placer. Appl Sci 10(2):635
16. Salem NM, Mahdi HMK, Abbas HM (2021) A novel face inpainting approach based on guided deep learning. In: 4th International conference on communications, signal processing and their applications
17. Salem NM, Mahdi HMK, Abbas HM (2019) Random-shaped image inpainting using dilated convolution. Int J Eng Adv Technol (IJEAT) 8(6)
18. Salem NM, Mahdi HM, Abbas H (2018) Semantic image inpainting using self-learning encoder-decoder and adversarial loss. In: 13th International conference on computer engineering and systems (ICCES). IEEE, pp 103–108
19. Sharma S, Ahmed S, Naseem M, Alnumay WS, Singh S (2021) A survey on applications of artificial intelligence for pre-parametric project cost and soil shear-strength estimation in construction and geotechnical engineering. Sensors 21:463. https://doi.org/10.3390/s21020463

Symptom Network Analysis of Social and Mental Health Complications of Alcohol Use Disorder

Kimasha Borah, Dhrubajyoti Chetia, Kalyan Bhuyan, and Dhrubajyoti Bhuyan

Abstract *Introduction* Alcohol use disorder is a global problem and is associated with various physical, mental and social complications. Symptom network analysis can provide an unique insight into the complex interaction of the various factors as regards to their contribution to all the social and mental health-related issues and vice versa. *Objective* This study attempts to understand this relationship in patients with alcohol use disorders by using principles of social network. *Methods* Relevant information was collected from patients attending Drug Deaddiction Centre of Assam Medical College, Dibrugarh, by using a structured tool containing questions related to pattern of use of alcohol and social and mental complications attributable to alcohol. *Results* The analysis of data by Gephi software showed that regular drinking had high degree score of 19 for social complications. In addition, fights with family members and decreased work performance (weighted degree 19 and 17, respectively) emerged as prominent contributors to social difficulties apart from regularity of drinking in the study population. Harmonic centrality values reveal that frequency of drinking behaviour is more influential than other factors in causing social harms. The degree centrality for network related to alcohol use and mental health-related complications was highest for regularity of drinking followed by stress and depression. *Conclusion* Frequency of drinking, fights with family members and decreased work performance are prominent contributors to social difficulties while regular drinking is a factor that is even more important than stress or depression in affecting mental health of persons with a history of alcohol.

K. Borah (✉)
Centre for Computer Science and Applications, Dibrugarh University, Dibrugarh, Assam, India
e-mail: kimasha08@gmail.com

D. Chetia
Lokapriya Gopinath Bordoloi Regional Institute of Mental Health, Tezpur, Assam, India

K. Bhuyan
Department of Physics, Dibrugarh University, Dibrugarh, Assam, India

D. Bhuyan
Assam Medical College, Dibrugarh, Assam, India

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_65

Keywords Alcoholism · Pattern of drinking · Betweenness centrality · Degree centrality · Eigen centrality

1 Introduction

Prolonged use of alcohol has been linked to a wide array of adverse health consequences spanning physical, mental and psychosocial domains. Current evidence suggests that the adverse effects of alcohol do not spare any human organ system. It is generally agreed that the link between alcohol and its adverse health consequences is dependent upon two prime factors. The first of these is the volume and pattern of drinking, and second equally important is the mediating mechanism of adverse health outcome, i.e. its biochemical effects, intoxication and dependence. Intoxication appears to be a mediator for acute outcomes by way of increased propensity to accidents, intentional or unintentional injuries, socio-familial conflicts and violence [1–3]. Notwithstanding the fact that dependence is itself a consequence of drinking, by virtue of its habit sustaining character alcohol dependence impacts both acute and chronic health events with serious adverse outcomes.

In the light of its serious and far reaching negative outcome, harmful use of alcohol can be regarded as a social malady. As per data released by the WHO in its Global status report on alcohol and health 2018, it is estimated that approximately 2.3 billion people are current drinkers. Although the number of users of alcohol has not increased in SE Asia, the per capita consumption volume has increased by 0.9 L of alcohol between 2005 and 2016. The same report also highlights the fact that as of 2016, 5.3% of all deaths is a consequence of harmful alcohol use, accounting for 5.1% of DALYs during that year. These figures are higher than those attributable to diseases such as diabetes, tuberculosis or HIV/AIDS. Within the wide array of adverse health consequences attributable to alcohol, 28.7% and 21.3% were due to injuries and digestive diseases, respectively. Acute adverse outcomes such as injuries accounted for as much as 40 pc of DALYs attributed to alcohol. A larger proportion (49%) of DALYs was on account of more subtle and chronic effects such as noncommunicable and mental health conditions [4, 5]. Also, of note is the fact that these disease-related burden due to alcohol is higher in LAMICs such as India.

Given that alcohol use stems from a complex interplay between bio-psycho-social factors, an enquiry into various aspects of alcohol use and its health consequences is in order. We intend to study this phenomenon from the perspective of network connectedness.

The current study is aimed at outlining the relationship between frequency of drinking and related health problems attributable to alcohol use using network analysis as a tool to define connectedness of aspects studied.

2 Materials and Methods

The sample was collected from patients attending Drug Deaddiction Centre of Assam Medical College, Dibrugarh. The data were collected from individuals with the diagnosis of alcohol use disorders in the age group of 18–60 years. Patients with comorbid or pre-existing mental illness or sub-normality and those who are critically ill to participate in the study or unwilling to participate were excluded from the study.

The relevant information was collected using a structured tool developed for the purpose of the study. The tool contains and covers questions relating to various aspects of alcohol use by the participants and outcomes attributable to alcohol in addition to sociodemographic data.

Data so obtained were entered into and analysed on the Gephi software. Gephi is an open-source network analysis and visualization software package written in Java on the NetBeans platform. Its latest version is Gephi 0.9.0 which was launched in December 2015, with subsequent updates in February 2016 (0.9.1) and September 2017 (0.9.2). It is fast emerging as a popular visualization and exploration software for all kinds of graphs and networks. It can be used for a wide range of applications such as Exploratory Data Analysis, Link Analysis, Social Network Analysis and Biological Network Analysis [6–9].

Two separate graphs were created from the dataset derived from the structured tool, the grouping being done based on the area of inquiry, viz social correlates of alcohol use and psychological consequences of alcohol use among the study participants. The layout algorithm was Fruchterman-Reinglod layout [7, 8], since it, among others if well suited to highlight complementaries in a dataset. These graphs were analysed for their network properties.

3 Results

3.1 Alcohol and Social Difficulties

Figure 1 is a depiction of the network resultant from graphical layout of the relevant variables in terms of social difficulties emerging out of alcohol use. Alcohol usage frequency among the study subjects was clustered into 4 groups, viz 1–2 days a week, 3–4 days a week, 5–6 days a weeks and daily/regular use. The graphical output contains 25 nodes with 47 edges. The graph has a network diameter of 5 and graph density being 0.157. The average degree is 3.76 (Avg weighted degree = 10.24). The average path length of the edges is 2.12.

The various measures of network properties of the output graph are depicted in Table 1.

Higher the degree and weighted degree, greater is the connectivity of a particular node to other nodes in the network. In our study, this connection may be interpreted as the extent of influence (directed or undirected) that the properties of the node

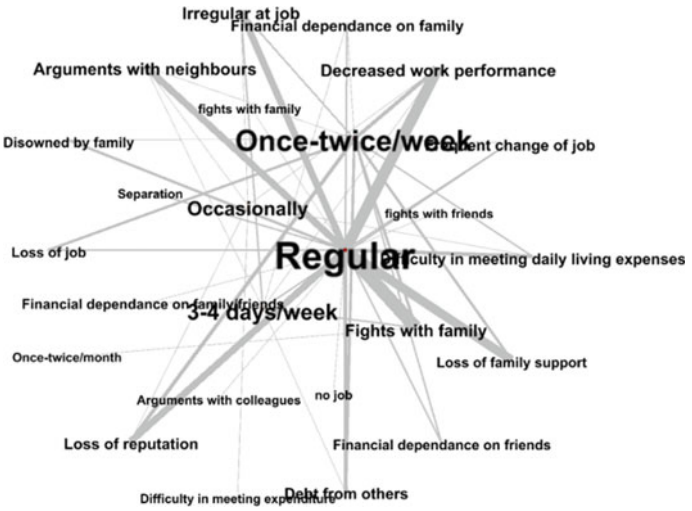


Fig. 1 Network of social difficulties

exert towards the functional output of the network studied viz towards causation of social difficulties and mental health problems, as well as being influenced by all other connected elements in the network. As is evident from the Table 1, the three most important factors that are related to social difficulties in the study population all relate to the frequency of drinking behaviour. These are regular drinking, drinking once or twice in a week and three/four times in a week as inferred from their high degree scores of 19, 13 and 8, respectively. In addition to the frequency of drinking, fights with family members and decreased work performance (weighted degree 19 and 17, respectively) emerged as prominent contributors to social difficulties in the study population. The components with the least weighted degree of 1 are having no job, arguments with colleagues, fights with friends, fights with family and difficulty meeting expanses.

The relative importance of a node in a network is further described with other centrality measures. Table 1 depicts the centrality measure values, and Figs. 2, 3, 4 and 5 depicts the various centrality distributions of the study population. The interpretation is undirected (diameter 5, radius 3 and average path length 2.126.)

As shown in Table 1, the three nodes representing the varied patterns of drinking frequency have the highest betweenness centrality. This was followed by arguments with neighbours and irregularity at work.

As was the case with degree centrality, regular drinking and drinking once or twice a week were the nodes with highest closeness centrality, followed by arguments, irregularity at work and decreased work performance. The node with the least closeness centrality value was difficulty meeting expanses. However, examination of the harmonic centrality values reveals that frequency of drinking behaviour is more influential than other factors. The nodes with the highest harmonic centrality

Table 1 Measures of network properties (social difficulties)

Id	Degree	Weighted degree	Eccentricity	Closeness centrality	Harmonic closeness centrality	Betweenness centrality	Authority	Hub	Pageranks	Eigen centrality
Occasionally	6	6	4	0.4138	0.5174	26.4167	0.2258	0.1039	0.0620	0.3596
Arguments with neighbours	4	10	3	0.5333	0.5764	10.1954	0.1312	0.2850	0.0399	0.4520
Loss of reputation	3	12	3	0.5106	0.5486	6.7000	0.1053	0.2288	0.0317	0.3636
Irregular at job	4	12	3	0.5333	0.5764	10.1954	0.1312	0.2850	0.0399	0.4520
Decreased work performance	4	17	3	0.5333	0.5764	10.1954	0.1312	0.2850	0.0399	0.4520
Difficulty in meeting expenditure	1	1	5	0.2963	0.3417	0.0000	0.0174	0.0378	0.0148	0.0611
Difficulty in meeting daily living expenses	3	7	3	0.5106	0.5486	6.7000	0.1053	0.2288	0.0317	0.3636
Regular use of alcohol	19	80	3	0.8000	0.8889	166.8517	0.6287	0.2893	0.1971	1.0000
Fights with family	4	19	4	0.5106	0.5660	25.2100	0.1171	0.2544	0.0442	0.4037
Debt from others	3	8	4	0.4898	0.5382	2.2100	0.1138	0.2473	0.0311	0.3909
Loss of family support	2	12	4	0.4706	0.5104	0.5115	0.0879	0.1910	0.0229	0.3024
Financial dependence on family	3	4	4	0.4898	0.5382	2.2100	0.1138	0.2473	0.0311	0.3909

(continued)

Table 1 (continued)

Id	Degree	Weighted degree	Eccentricity	Closeness centrality	Harmonic closeness centrality	Betweenness centrality	Authority	Hub	Pageranks	Eigen centrality
Drinking 3–4 days/week	8	11	3	0.4615	0.5833	16.1577	0.3365	0.1548	0.0771	0.5316
Drinking once-twice/week	13	30	3	0.5714	0.7222	49.5739	0.5138	0.2364	0.1243	0.8120
Financial dependence on friends	2	4	4	0.4706	0.5104	0.5115	0.0879	0.1910	0.0229	0.3024
Frequent change of job	3	5	4	0.4898	0.5382	2.2100	0.1138	0.2473	0.0311	0.3909
Loss of job	2	5	4	0.4706	0.5104	0.5115	0.0879	0.1910	0.0229	0.3024
No job	1	1	4	0.4528	0.4826	0.0000	0.0484	0.1051	0.0148	0.1679
Disowned by family	2	4	4	0.4706	0.5104	0.5115	0.0879	0.1910	0.0229	0.3024
Arguments with colleagues	1	1	4	0.4528	0.4826	0.0000	0.0484	0.1051	0.0148	0.1679
Fights with friends	1	1	4	0.4528	0.4826	0.0000	0.0484	0.1051	0.0148	0.1679
Drinking once-twice/month	1	1	5	0.3429	0.3729	0.0000	0.0425	0.0196	0.0154	0.0678
Separation	1	2	4	0.4528	0.4826	0.0000	0.0484	0.1051	0.0148	0.1679
Financial dependence on family/friends	2	2	4	0.4706	0.5104	1.1278	0.0743	0.1614	0.0230	0.2563
Fights with family	1	1	4	0.4528	0.4826	0.0000	0.0484	0.1051	0.0148	0.1679

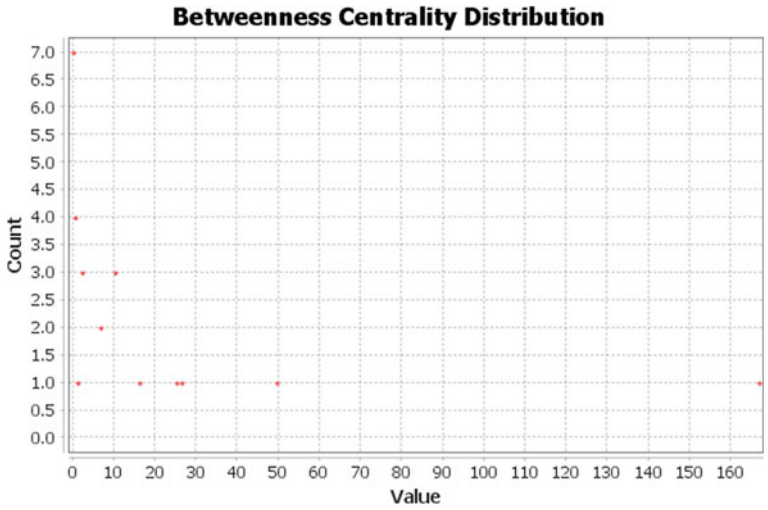


Fig. 2 Betweenness centrality distribution of network properties (Social difficulties)

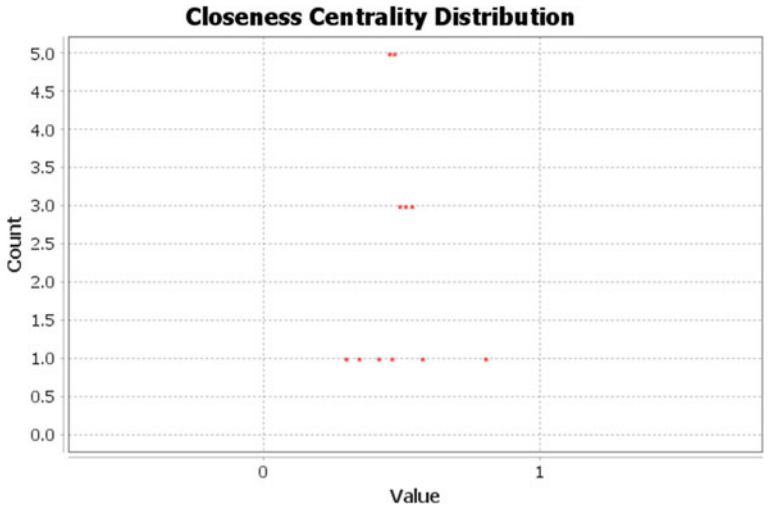


Fig. 3 Closeness centrality distribution of netwok properties (Social difficulties)

are regular drinking, drinking once or twice a week and drinking 3/4 times a week followed by arguments with neighbours and irregularity at job.

The eigen centrality values of our dataset are depicted in Table 1 and distribution in Fig. 3, respectively (Number of iteration 1000; Sum change 4.309). The eigen centrality values are a measure of how influential a particular variable is towards affecting the overall functional outcome of the network node interactions—in our case the outcome being the domain of social difficulties. The three varied patterns

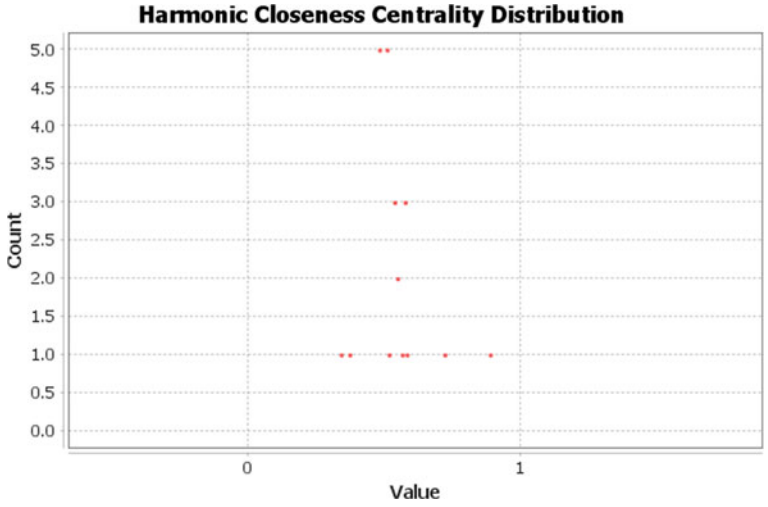


Fig. 4 Harmonic closenness centrality distribution of network properties (Social difficulties)

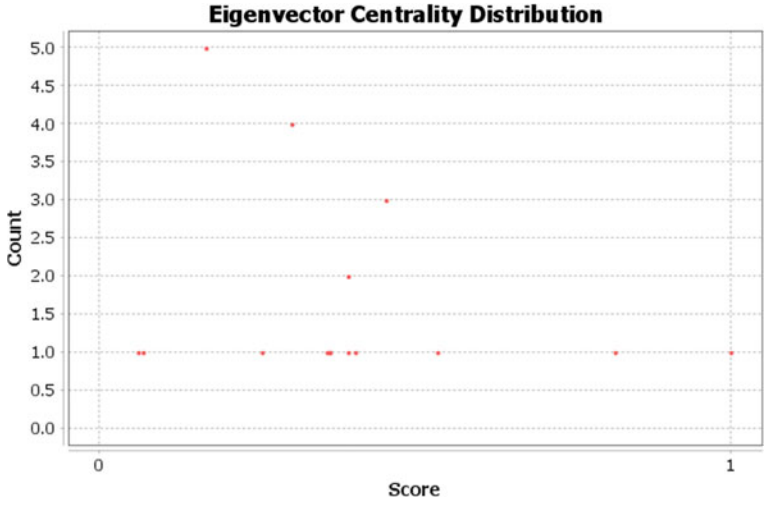


Fig. 5 Eigenvector centrality distribution of network properties (Social difficulties)

of drinking frequency again emerge as the nodes with the highest eigen centrality values, implying that frequency of drinking behaviour is central to overall expression of social difficulties that persons with alcohol use face. The other important nodes as reflected by eigen centrality scores are arguments with neighbours and irregularity at job. The least influential factor was difficulty maintaining expanses.

4 Alcohol and Mental Health Issues

The second graph we created took into account the frequency of alcohol use and the experienced mental health problems of the study population. The Fruchterman–Reingold algorithm was used. The network so derived is depicted in Fig. 6. The undirected graph contains 13 nodes and 19 edges. The graph has a network diameter of 4 and graph density being 0.244. The average degree is 2.923 (Avg weighted degree = 6.46). The average path length of the edges is 2.154.

The various parameters of this network are detailed in Table 2.

The distribution of network parameters is depicted in Figs. 7, 8, 9 and 10

The degree centrality for this network was highest for regularity of drinking followed by stress and depression. As stated earlier, degree centrality of a node is a measure of the extent to which the node id connected to other nodes. We can thus infer that regular drinking is a factor that is even more important than stress or depression in affecting mental health of persons with a history of alcohol use (without implying causality). The factors with the least degree centrality were having undefined problems and associated guilt. On weighted degree measures although regular drinking remained the factor with the highest value, depression was shown to be more influential than stress per se in affecting mental health. Also, on weighted degree measures, feeling guilty appears to be a more important factor than undefined problems although on degree centrality measures; both appears to be at par in affecting mental health.

The factor with the highest betweenness centrality is again regularity of drinking, followed by stress and feeling depressed. This can be interpreted to mean the involvement of these factors as being the prime factors that mediate the expression of mental health outcome of the other nodes in the network. Viewed as such an indicator, it implies that the overt manifestation of a mental health issue and its relation to another variable is mediated primarily by frequency of drinking and the presence or absence

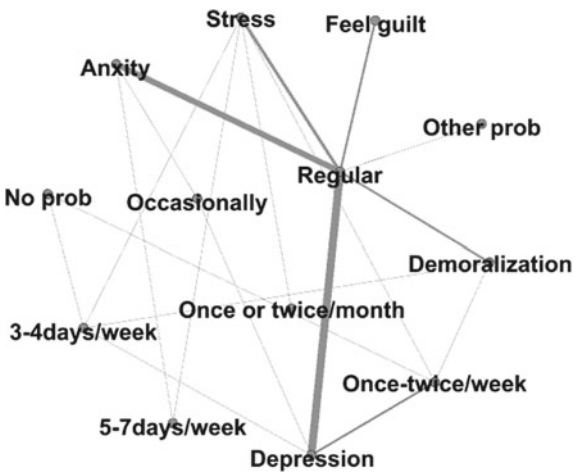


Fig. 6 Network of mental health-related issues

Table 2 Measures of network properties (mental health related issues)

Id	Degree	Weighted degree	Eccentricity	Closeness centrality	Harmonic closeness centrality	Betweenness centrality	Pageranks	Eigen centrality
Occasionally	2	2.000	4.000	0.414	0.493	1.300	0.056	0.377
Depression	4	14.000	3.000	0.545	0.639	10.900	0.099	0.849
Anxiety	3	9.000	4.000	0.462	0.563	5.700	0.081	0.498
Regular use of alcohol	6	27.000	3.000	0.632	0.736	29.167	0.156	1.000
Stress	5	8.000	3.000	0.600	0.694	21.300	0.127	0.928
Drinking 3-4 days/week	4	4.000	3.000	0.522	0.625	7.767	0.098	0.832
Drinking once-twice/week	4	6.000	3.000	0.522	0.625	7.767	0.098	0.832
Demoralization	3	5.000	3.000	0.500	0.583	3.850	0.075	0.743
No problem	2	2.000	4.000	0.375	0.465	0.250	0.053	0.463
Other problems	1	1.000	4.000	0.400	0.451	0.000	0.034	0.280
Feel guilt	1	3.000	4.000	0.400	0.451	0.000	0.34	0.280
Drinking 5-7 days/week	2	2.000	3.000	0.444	0.514	2.000	0.056	0.399
Drinking once-twice/month	1	1.000	4.000	0.387	0.438	0.000	0.033	0.259

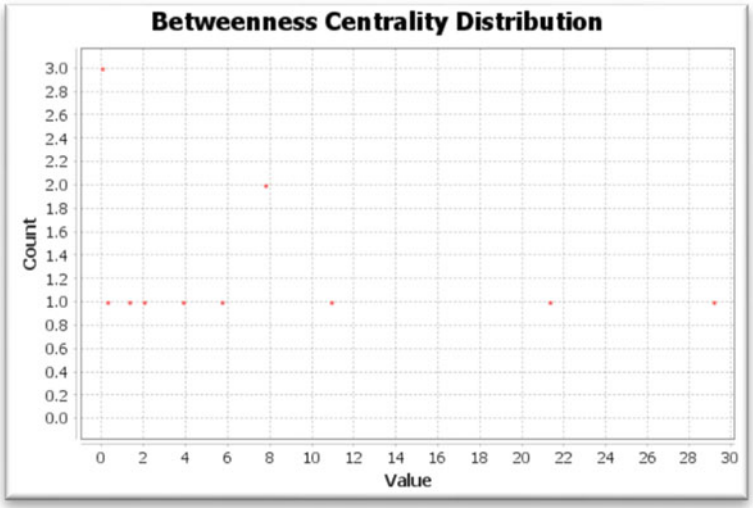


Fig. 7 Betweenness centrality distribution of network properties (mental health related issues)

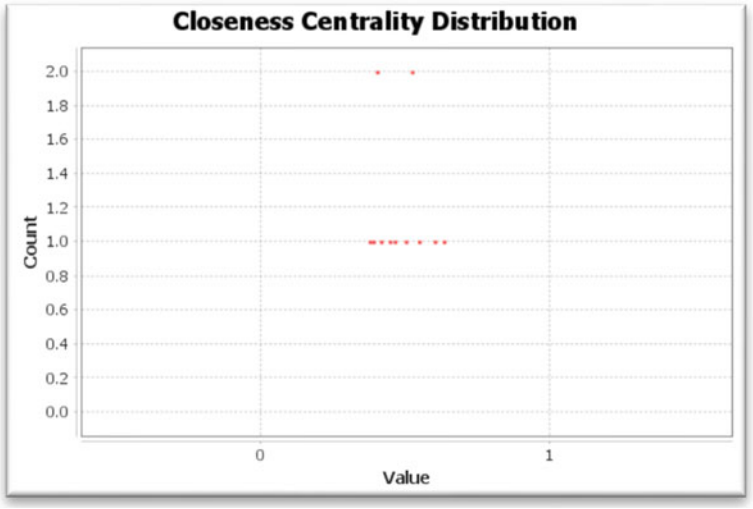


Fig. 8 Closeness centrality distribution of network properties (mental health related issues)

of stress and feeling depressed. For example, in our study population, there is no direct link between anxiety and stress; however, the two are connected by high weight edges passing through the node representing regular drinking, in addition to other pathways with less edge weight.

The inference from examination of harmonic closeness centrality values is parallel to the inference from closeness centrality and reflects the importance of frequency

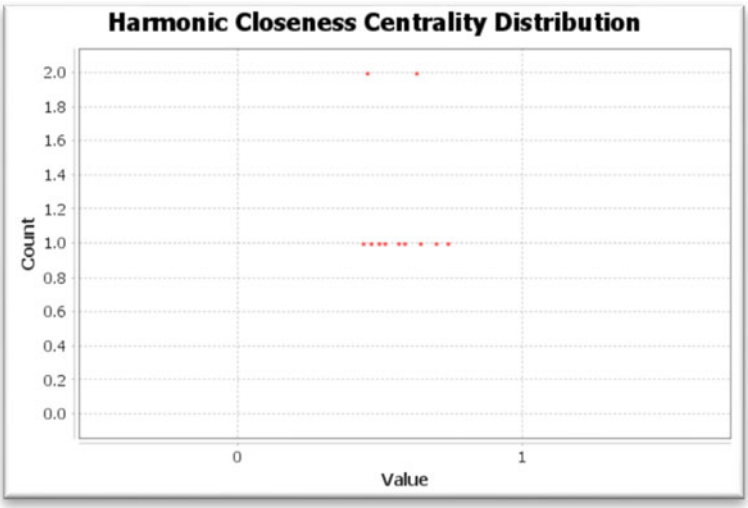


Fig. 9 Harmonic Closeness centrality distribution of network properties (mental health related issues)

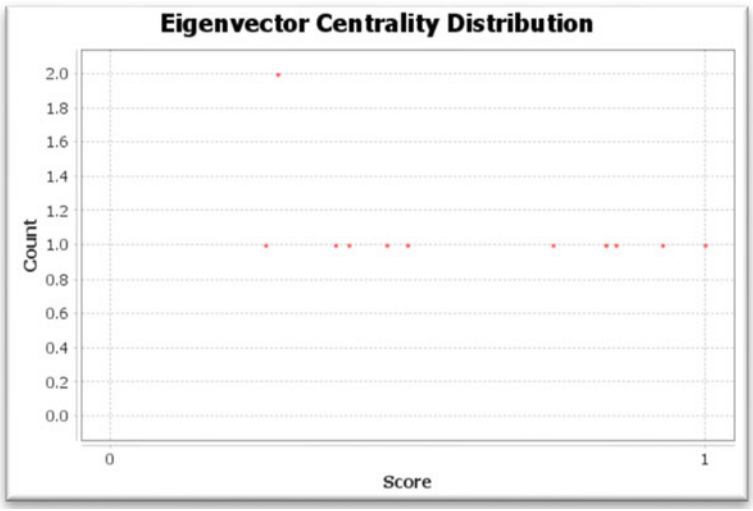


Fig. 10 Eigenvector centrality distribution of network properties (mental health related issues)

of drinking, stress and feeling depressed as important components of mental health outcomes in the study population.

A similar pattern of influence is revealed by eigenvector centrality scores. As stressed before, eigen vector or prestige is a measure of the influence of an individual node in a network. In other words, eigen vector centrality scores can identify which

nodes play a pivotal role in determining the functional outcome of a network. The pattern that we see here runs parallel to other centrality values, i.e. regular drinking, stress and feeling depressed are the three most influential factors the determine mental health outcome of persons using alcohol.

5 Conclusion

Most of the researches on alcohol use disorder focus on collection and analysis of data in terms of criteria laid down by the nosological systems i.e. ICD10 or DSM 5. The social, occupational, physical and other related harms have been the topic of interest in various literatures and also in various types of psychosocial interventions directed towards ensuring abstinence. Network theory provides an unique opportunity to look into the relationship of various factors related to drinking habit and pattern with the associated harms. In other words, network model allows the symptom-symptom interaction which might help the clinicians in understanding this dynamics in a simplistic way.

This report focuses on two key issues:

1. Alcohol use and social difficulties: Examination of the harmonic centrality values reveals that frequency of drinking behaviour is more influential than other factors.
2. Alcohol use and mental health issues: Frequency of drinking, stress and feeling depressed as important components of mental health outcomes in the study population.

Contrary to the popular believe, inability to maintain the expenses has been found to be least important in causation of social difficulties and frequency of alcohol use has been found to be the root cause of the social complications. Therapies targeted to reduce the frequency of alcohol use and management of stress and depression can be expected to be of key importance in reducing the mental health-related issues.

References

1. Rehm J (2011) The risks associated with alcohol use and alcoholism. Alcohol Res Health: J National Institute on Alcohol Abuse Alcoholism 34(2):135–143
2. Sarkar D, Jung MK, Wang HJ (2015) Alcohol and the immune system. Alcohol Res: Current Rev 37(2):153–155
3. Schulte MT, Hser YI (2014) Substance use and associated health conditions throughout the lifespan. Public Health Rev 35(2). https://web-beta.archive.org/web/20150206061220/http://www.publichealthreviews.eu/upload/pdf_files/14/00_Schulte_Hser.pdf. https://doi.org/10.1007/BF03391702
4. World Health Organization (2018) Global status report on alcohol and health 2018

5. GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2018; 392: 1015–35
6. Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. In: Third international AAAI conference on weblogs and social media
7. https://gephi.org/tutorials/gephi-tutorial-layouts.pdf
8. Hawe P, Webster C, Shiell A (2005) A glossary of terms for navigating the field of social network analysis. J Epidemiol Community Health 58:971–975. https://doi.org/10.1136/jech.2003.014530
9. Anand B, Manoj P (2015) Eigenvector centrality and its application in research professionals' relationship network. In: 2015 1st International conference on futuristic trends in computational analysis and knowledge management, ABLAZE 2015. https://doi.org/10.1109/ABLAZE.2015.7154915

Measuring E-government System Users' Satisfaction Using a Multicriteria Analysis Model: A Case Study of Botswana

Ezekiel U. Okike and Omphemetse N. Small

Abstract This paper presents a rigorous evaluation of Botswana's e-government system users' satisfaction using a multicriteria satisfaction analysis (MUSA) model. While most experiments found the MUSA model rigorous and very useful, one study suggested that the model does not always give an interpretable result. The present study evaluates the usefulness of MUSA in the analysis of users' satisfaction of Botswana e-government systems with a sample of 136 users comprising 35 IT professionals and 101 non-IT professionals. This sample was selected using a purposive sampling to select employees who have used at least one e-government system, and focusing on Botswana government ministries who first took up e-government systems initiatives. The results indicate that MUSA was appropriate in evaluating the system, and it gives interpretable results. Also, Botswana e-government systems has a global satisfaction index of 0.86 (86%) rating, a partial satisfaction index of 0.5447 (54.5%), a demand index of 0.385 (38.5%). The partial satisfaction index measurements for individual system components are actual use (44%), usability (81%), functionality (19%), efficiency (51%), reliability (27%), maintainability (29%), portability (41%), and operational readiness (25%). The sampled users are non-demanding users and are satisfied with the systems. Based on the partial satisfaction indices, components of the system with low ratings need improvement.

Keywords Measurement · E-government systems · Users' Satisfaction · Multicriteria Analysis · Global Satisfaction

E. U. Okike (✉) · O. N. Small
University of Botswana, Gaborone, Botswana
e-mail: okikeue@ub.ac.bw

O. N. Small
e-mail: smalln@ub.ac.bw

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_66

1 Introduction

Measurement governs every aspect of life including determining prices using economic indicators, diagnosing diseases and illnesses in medicine, and predicting the weather in meteorology. The process of measurement assigns identifiers to attributes of entities of interest in the real world such that it describes properly such entities according to defined rules [1]. In software systems, rigorous measurements are used to determine software quality by applying available well-established models or metrics such as Lines of Code Metric, McCabe Cyclomatic Number Metric, Halstead Software Science Metrics, Maintainability Index Metric, Software Reliability Metrics, and other Programming Complexity Measures [2–6]. The need and usefulness of measurements in every discipline of study cannot be over-emphasized. As in the well-established science and engineering disciplines, all measurement must follow a rigorous approach and conform with the rules of measurement theory [1, 7, 8]. As countries and nations of the world embrace the application of ICT in governance and using the same to enhance delivery of services, e-government system success, usefulness, and users' satisfaction must be measured with appropriate and reliable models to ascertain their usefulness, service delivery, and overall readiness.

As an information system (IS), e-government systems have recently evolved with the aim to effectively and efficiently deliver government services to the people. As such, national governments following their own defined information system strategy (ISS) implement e-government systems according to their needs. The ISS defines the organization requirement or demand for information and systems to support the overall strategy of the business (in this case e-government business). Therefore, e-government systems are anchored on the overall information system strategy. The strategy according to [9] "defines and prioritizes the investments required to achieve the ideal application portfolio, the nature of the benefits to expect, and the changes required to deliver these benefits within the constraints of resources and systems interdependence."

1.1 Statement of the Problem

There is a current need for a useful and rigorous satisfaction model to measure the satisfaction of users of e-government systems. While most experiments found the MUSA model rigorous and very useful, one study suggested that the model does not always give an interpretable result [10]. This needs further investigation. Moreover, while most African countries have already deployed e-government systems, there is no evidence from the literature that any of the systems have been evaluated using rigorous and reliable models to determine the satisfaction of users of the systems.

1.2 Study Objectives

This study evaluates users' satisfaction of Botswana's e-government systems by means of a multicriteria satisfaction analysis (MUSA) model to determine if the MUSA model is useful in the evaluation of e-government systems with interpretable results.

2 Literature Review and Model Explanation

In Africa, the development and deployment of e-government systems have been well documented in the literature [11–17]. Most countries have functional e-government systems at various levels of deployment status in government ministries for very essential services. It is hoped that with steady progress every government-based service would be covered appropriately. This being the case the need for appropriate user (customer) satisfaction evaluation of the systems cannot be over-emphasized. Some of the advantages of such evaluation include:

(a) Improving communication with the users.
(b) Assisting the organization in examining whether the provided services fulfill users' expectations.
(c) Assisting the organization in determining critical dimensions that needs improvement.
(d) Such evaluations might be the source of motivation to increased staff productivity since all improvement efforts regarding the offered service are evaluated by the users themselves.

We define user (customer) satisfaction in the context of this study according to [18] to mean the "fulfillment response, including the judgment that a product or service feature, or product or service itself provided (or is providing) a pleasurable level of consumption related fulfillment including level of under or over fulfillment." While a number of works such as [10, 18–21] applied different approaches to measure user (customer) satisfaction, the present study is based on a multicriteria analysis model (MUSA).

2.1 The Musa Model

The model as presented in [22–24] is formerly defined as follows:

$$\widetilde{Y}^* = \sum_{i=1}^{n} b_i X_i^* - \sigma^+ + \sigma^- \tag{1}$$

With

$$\sum_{i=1}^{n} b_i = 1$$

where $\widetilde{Y}^*$ is the estimation of the global value function Y^*, X_i* is an estimation of the partial satisfaction**,** **n** is the number of criteria, $\mathbf{b_i}$ is a positive weight of the *i*th criterion, σ^+ and σ^- are the overestimation and the under-estimation errors, respectively, and the value functions are normalized in the interval [0,100].

The methodology also allows the formulation of linear programming model using goal seeking techniques to optimize user satisfaction measurement. The optimization model is depicted in Eq. (2)

$$\begin{cases} [\min] F = \sum_{j=1}^{M} \left(\sigma_j^+ + \sigma_j^-\right) \\ \text{under the constraints} \\ \sum_{i=1}^{n} \sum_{k=1}^{t_{ii}-1} w_{ik} - \sum_{m=1}^{t_j-1} z_m - \sigma_j^+ + \sigma_j^- = 0 \quad \text{for } j = 1, 2, \ldots, M \\ \sum_{m=1}^{\alpha-1} z_m = 100 \\ \sum_{i=1}^{n} \sum_{k=1}^{\alpha-1} w_{ik} = 100 \\ z_m \geq 0,\ w_{ik} \geq 0 \quad \forall m, i, k \\ \sigma_j^+ \geq 0,\ \sigma_j^- \geq 0 \quad \text{for } j = 1, 2, \ldots, M \end{cases} \tag{2}$$

where M is the size of the user's sample, while y^j and $x_i{}^j$ are the *j*th levels on which variables Y and X_i are estimated (i.e., global and partial satisfaction judgments of the *j*th user).

To reduce mathematical computations, transformation Eq. (3) is applied according to [19, 22–24]:

$$\begin{cases} z_m = y^{*m+1} - y^{*m} & \text{for } m = 1, 2, \ldots, \alpha - 1 \\ w_{ik} = b_i x_i^{*k+1} - b_i x_i^{*k} & \text{for } k = 1, 2, \ldots, \alpha_i - 1 \text{ and } i = 1, 2, \ldots, n \end{cases} \tag{3}$$

where y^{*m} is the value of the y^m satisfaction level, x_i^{*k} is the value of the x_i^k satisfaction level, and α is the number of global and partial satisfaction levels.

2.2 Global Satisfaction Index

The global satisfaction index is the average of overall satisfaction that the users attach to a specific criterion. Because of this, it is often considered as the basic average performance indicator, range [0,100] for the business organization or government. Equation 4 represents a model of global satisfaction.

$$S = \frac{1}{100} \sum_{m=1}^{\alpha} p^m y^{*m} \tag{4}$$

where S is the average global satisfaction index, and p^m are the frequencies of users belonging to the y^{*m} satisfaction levels.

2.3 Average Demanding Indices

Global and partial demanding indices, D and D_i, respectively, are assessed as follows:

$$\begin{cases} D = \dfrac{\sum_{m=1}^{\alpha-1} \left(\frac{100(m-1)}{\alpha-1} - y^{*m} \right)}{100 \sum_{m=1}^{\alpha-1} \frac{m-1}{\alpha-1}} & \text{for } \alpha > 2 \\ D_i = \dfrac{\sum_{k=1}^{\alpha_i-1} \left(\frac{100(k-1)}{\alpha_i-1} - x_i^{*k} \right)}{100 \sum_{k=1}^{\alpha_i-1} \frac{k-1}{\alpha_i-1}} & \text{for } \alpha_i > 2 \text{ and } i = 1, 2, \ldots, n \end{cases} \tag{5}$$

It should be noted that:

$D = 1$ or $D_i = 1$—Users have the highest demanding index.
$D = 0$ or $D_i = 0$—This case refers to "neutral" users.
$D = -1$ or $D_i = -1$—Users have the lowest demanding index.

Moreover, MUSA methodology also provides a series of normalized indices that may help in-depth analysis of the satisfaction measurement problem.

The partial satisfaction index (Eq. 6) shows the level of partial satisfaction of the users (S_i) to every criterion, sub-criterion, and overall satisfaction index.

$$S_i = \frac{1}{100} \sum_{k=1}^{\alpha_i} p_i^k x_i^{*k} \quad \text{for } i = 1, 2, \ldots, n \tag{6}$$

In this study, user's global or overall satisfaction depends on eight top level criteria as shown in satisfaction evaluation hierarchy Fig. 1. The hierarchy also shows that each of the eight criteria has sub-criteria. The sub-criterions form the partial satisfaction dimensions. Responders were asked to express their opinion on a five-point Likert scale (1–5) with point 1 and 5 being the lowest and highest scales, respectively.

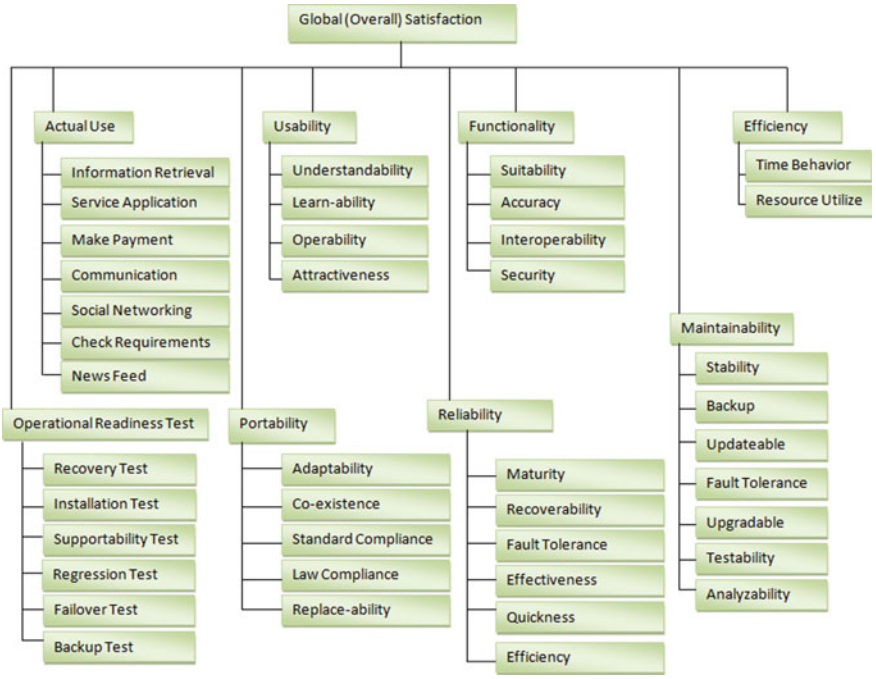


Fig. 1 Satisfaction evaluation criteria hierarchy

2.4 Justification of the Choice of Musa Method

First, most frameworks found in literature for assessing e-government and e-government user satisfaction focus on performance metrics or dimensions and leave out the assignment of weights or importance indicating how those metrics will be measured. To fill up this gap, the MUSA methodology was used since it enables users to assign weights to individual criterions/dimensions used in the evaluation process.

Second, studies that adopted ISO/IEC 9126 considered the metrics at high levels, meaning that they used the high-level evaluation criteria as a whole without considering the sub-metrics under each criterion. The MUSA method enables hierarchical weighted assessment considering both major criteria (high level or global metrics) and sub-criteria (low level or partial metrics).

3 Research Method

This study adopted a quantitative research method, using a survey study approach. The survey instrument was a questionnaire designed to elicit appropriate information from users of any e-government system in Botswana specifically asking for how

Table 1 Research variables/constructs

Dimension	Source
System Functionality	ISO/IEC 9126; ISO/IEC 9126-1, 2001
System Usability	ISO/IEC 9126; ISO/IEC 9126-1, 2001
System Reliability	ISO/IEC 9126; ISO/IEC 9126-1, 2001
System Maintainability	ISO/IEC 9126; ISO/IEC 9126-1, 2001
System Portability	ISO/IEC 9126; ISO/IEC 9126-1, 2001
System Actual Use	E-GOVSAT
Operational Readiness Test	E-GOVSAT

satisfied the users were with the system in terms of functionality, usability, reliability, system maintainability, portability, actual use, and operational readiness. These evaluation criteria were adopted for the e-government systems from the ISO/IEC 9126 [25] standard base model for information systems. The questionnaire used a 5-point Likert scale ranging from "strongly disagree" with a rating of 1 to "strongly agree" with a rating of 5. Participants in the study were purposefully selected to include those who are computer literate, some IT professionals, and a participant must have used at least one Botswana e-government system.

A total of 170 questionnaires were administered to users of Botswana's e-government systems for a period of two weeks in July 2019. 136 responses were properly completed while 34 questionnaires were discarded due to incompleteness. Participation on the research was voluntary, and no rewards of any sort were attached to it. Table 1 below shows the variables used in evaluating the e-government system satisfaction from the ISO/IEC 9126 software quality model.

Furthermore, the questionnaire in this study was piloted with 5 users each of whom have used different e-government systems, namely the e-government systems for Immigration Citizenship System (ICS), Botswana Animal Information Traceability System (BAITS), District Health Information System (DHIS), Ministry of Trade and Industry Management Information System (MTIMIS), and Ministry of Land, Government, and Rural Development System (MLGRDS). The purpose of the pilot study was to ensure that the questionnaire adaptation to the current study was carefully worded and that it was understandable to the intended respondents and gives the anticipated responses. The result of the pilot study revealed that inclusion of employees in the informal sector as the study respondents might skew the result given their lack of experience and computer illiteracy with regards to e-government systems; thus, they were not to be included in the final study. Data collected for the study was analyzed using the Statistical Package for the Social Sciences (SPSS) [26].

3.1 Satisfaction Evaluation Criteria

A hierarchy of eight satisfaction criteria and their related sub-criteria used in this study is shown in Fig. 1.

3.2 Justification of Sample Size

Studies have shown that the sample size should aim at surpassing the 1.7% of the entire population in order to have a representative sample [27]. According to Statistics Botswana [28] (formal sector employment: first quarter report 2019), there are about ten thousand government employees in the formal sector in gaborone; a minimum of 170 users was considered adequate for the study [29]. In terms of survey instrument reliability, Table 2 shows Cronbach's values for individual constructs of the chosen satisfaction evaluation criteria.

As suggested in [29, 30], values of Cronbach's alpha for strength of measurement of constructs are as follows: excellent (≥0.9), very good (0.8 to <0.9), good (0.7 to <0.8), moderate (0.6 to <0.7), and poor (<0.6). Construct system functionality, system reliability, system maintainability, system usability, and operational readiness test all have Cronbach's alpha higher than 0.7 which indicates that in these constructs there is very good internal consistency.

4 Result and Discussion

The findings as presented in Table 3 below show the rate of return on research instruments.

Table 2 Cronbach's alpha table

Construct	Cronbach's alpha
System Actual Use	0.653
System Functionality	0.885
System Efficiency	0.410
System Reliability	0.872
System Maintainability	0.896
System Portability	0.659
System Usability	0.811
Operational Readiness Test	0.788
Overall	0.892

Table 3 Rate of return of questionnaires

	Frequency	Percentage (%)
Response	136	80
Non-response	34	20
Total	170	100

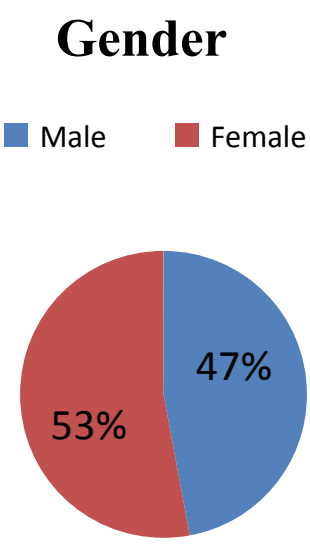
The results as presented in Table 3 above show that the study obtained a response rate of 80% (136), whereas 20% (34) did not respond. This adheres to [31] recommendations of 75% achievement for statistically viable data to make a conclusion and recommendation about a subject under study.

4.1 Demographic Information

This comprised of gender (Fig. 2), age group (Fig. 3), position (Table 4), and department (Fig. 4).

In all, there were 35 IT professionals and 101 non-IT professionals who participated in the study as shown in Table 5.

Fig. 2 Gender distribution



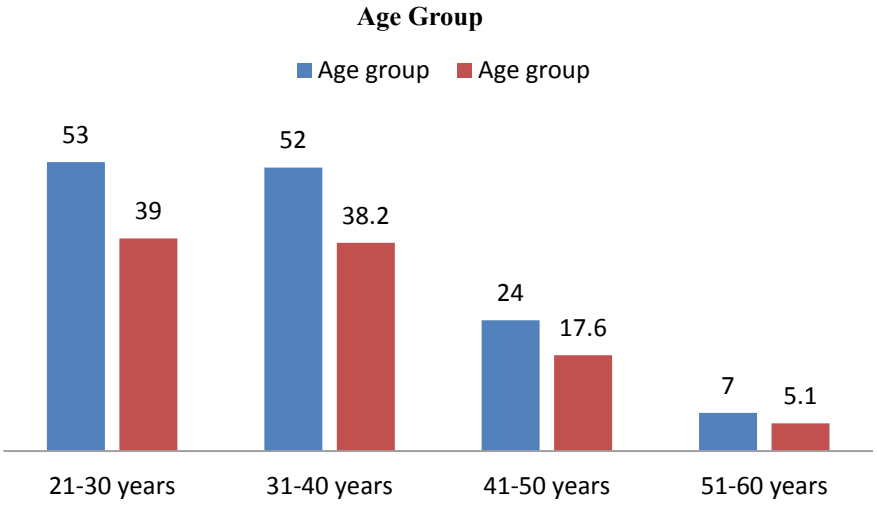


Fig. 3 Age groups

Table 4 Respondents' position categories

Administration	Public Relations	Public Servants	Marketing and Management	Information Technology	Farmer
Typist	Communications Officer	Nurse	Marketing Officer	IT Officer	Farmer
Receptionist	Communications Manager	Action Officer	Procurement Officer	System Analyst	
Administration Officer	Licensing Officer	Principal Registered Nurse	Management Analyst	Senior System Analyst	
Senior Administration Officer	Licensing Manager	Cleaner	Accountant	Chief Senior System Analyst	
Documentation and Distribution Officer	Public Relations Officer			Network Administrator	
Photographer					

4.2 Use of Botswana E-government Systems

The findings as presented in Table 6 below show the different e-government systems that the respondents to the study used in their respective working environments. The most used e-government systems were the Botswana Unified Revenue Service System 36 (26.5%), Botswana Post-Vehicle License Renewal System 28 (20.6%), and GABS 23 (16.9%).

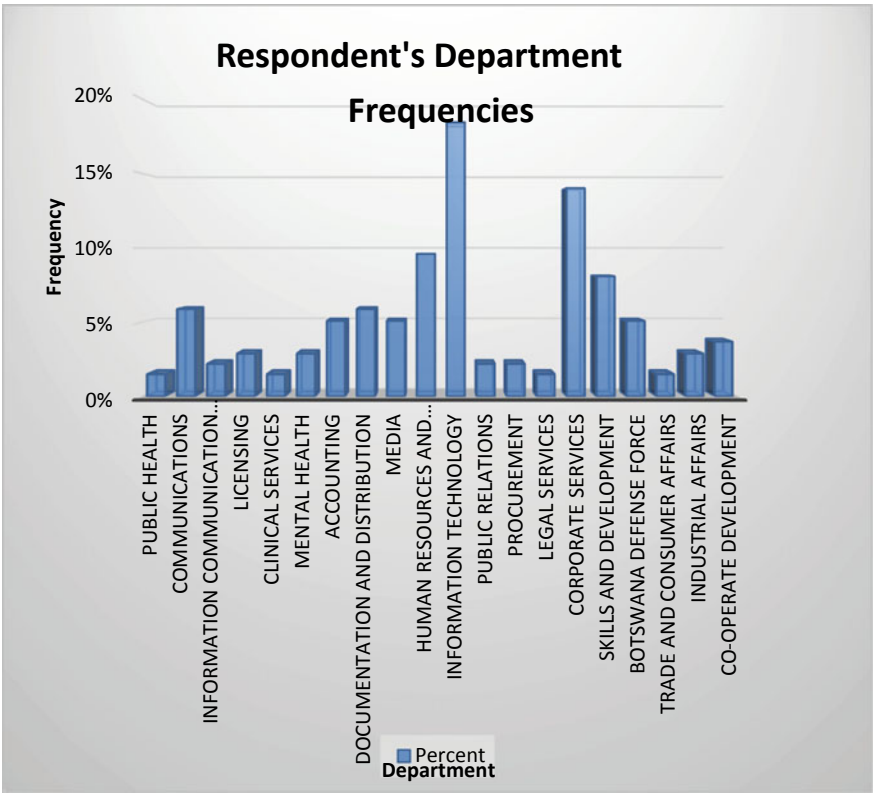


Fig. 4 Departments by categories

Table 5 Distribution of IT versus non-IT respondents

	Frequency	Percentage (%)
IT-related profession	35	25.7
Total non-IT-related professionals	101	74.3
Total	136	100

The results as presented in Fig. 5 below show participant response on how they got their system awareness.

4.3 User Satisfaction with E-government Systems

The result of applying a multicriteria satisfaction analysis (MUSA) model using 8 criteria (actual use, system usability, functionality, efficiency, reliability, maintainability, portability, and operational acceptance/readiness testing) is shown below.

Table 6 E-government systems

E-government systems		
	Frequency	Percent
Ministry of Investment, Trade, and Industry Management System	9	6.6
Botswana Post-Vehicle License Renewal System	28	20.6
Botswana Unified Revenue Service System	36	26.5
MLGRDS	9	6.6
Infinium	9	6.6
GABS	23	16.9
District Health Information System	2	1.5
Integrated Procurement Management Information	3	2.2
Botswana Animal Information Tracking System	6	4.4
Intsup	6	4.4
Immigration and Citizenship System	5	3.7
Total	136	10.0

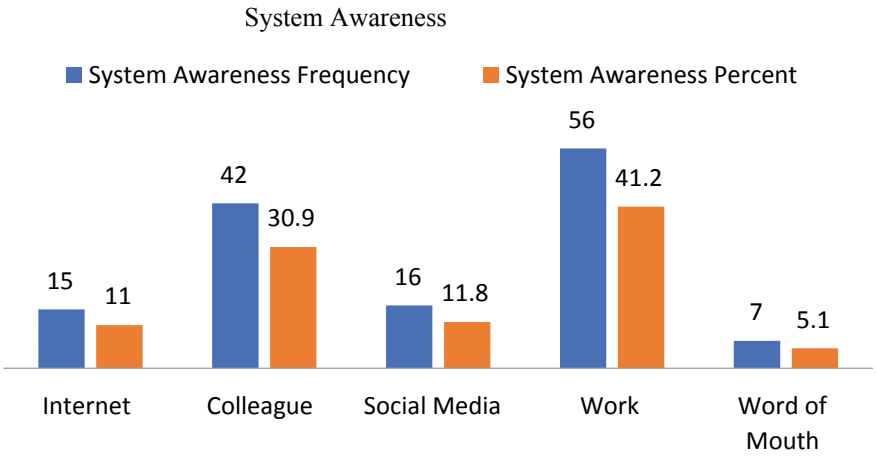


Fig. 5 System awareness

Aspects of maintainability, portability, and operational readiness tests were only answered by information technology professionals. Table 7 shows the function value increment of 25% for each level of the Likert scale.

Table 7 Function value increment

Level	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	0	25	50	75	100

Table 8 Linear programming criteria weights

	W_{ik}				b_i
	1	2	3	4	Weight
Actual Use	2.8	5.0	6.8	5.4	$(2.8 + 5.0 + 6.8 + 5.4)/100 = 20\%$
Usability	0.9	2.5	2.0	0.7	$(0.9 + 2.5 + 2.0 + 0.7)/100 = 6\%$
Functionality	2.5	5.2	4.2	0.2	$(2.5 + 5.2 + 4.2 + 0.2)/100 = 12\%$
Efficiency	2.7	1.9	8.5	0.9	$(2.7 + 1.9 + 8.5 + 0.9)/100 = 23\%$
Reliability	0.7	3.5	4.1	0.7	$(0.7 + 3.5 + 4.1 + 0.7)/100 = 9\%$
Maintainability	0.1	0.6	0.7	0.6	$(0.1 + 0.6 + 0.7 + 0.6)/100 = 2\%$
Portability	0.0	7.6	6.8	7.6	$(0.0 + 7.6 + 6.8 + 7.6)/100 = 22\%$
ORT	0.4	2.7	2.9	0.0	$(0.4 + 2.7 + 2.9 + 0.0)/100 = 6\%$

The function value (FV) for global satisfaction is estimated to be consistent with the summation of the function value (FV) for partial satisfaction. Any differences were solved with the error estimation.

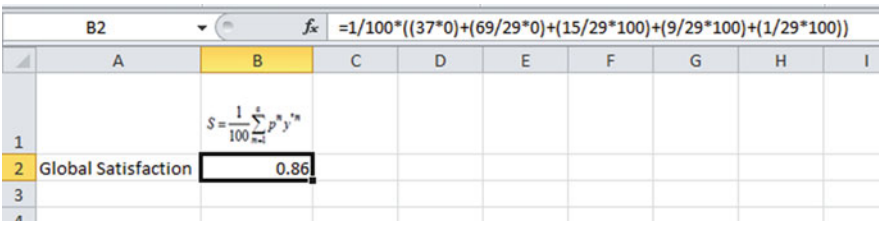
4.4 Linear Programming for Criteria Weights

Using linear programming (Eq. 2), the result for criteria weights was obtained as shown in Table 8 [17].

4.5 Global Satisfaction

The global satisfaction function is an important outcome of the MUSA model. In general, it appears users are quite satisfied with the provided service, given that the average global satisfaction index has quite high value (86%). Equation 4 was used as a model to generate global satisfaction as in Table 9. The global satisfaction demand curve is shown in Fig. 6.

Table 9 Global satisfaction index



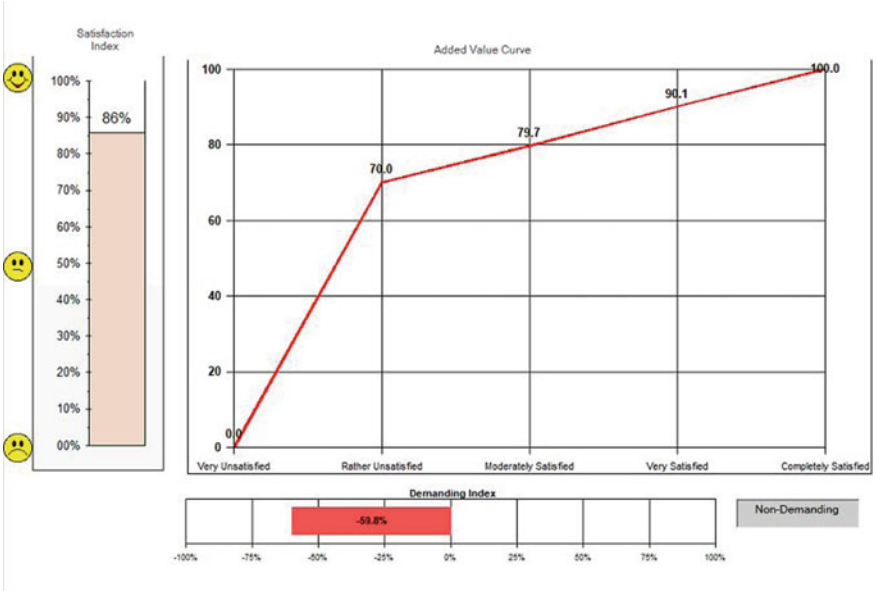


Fig. 6 Global satisfaction demanding curve

4.6 Satisfaction Indices

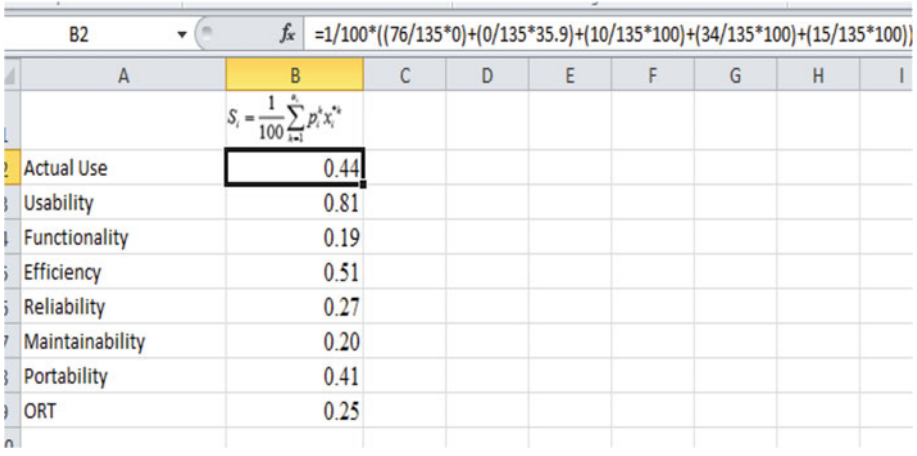
It is evident from Fig. 6 that the sample consists of non-demanding users. This is also confirmed by the 59.8% demanding index on the graph which falls under the non-demanding category of the ordinal demanding scale. Furthermore, the graph also reveals that respondents who are neutrally satisfied according to the value function are 79.7% which is relatively high which may mean that we deal with extremely non-demanding users or users who do not attach significant importance to the e-government systems (Fig. 6).

Using Eq. 6, partial satisfaction indices were calculated as follows for all 8 criteria (see Table 10).

The results as shown in Fig. 8 confirm that the users of e-government system services were mostly satisfied with its usability (81.30%). Other satisfaction results are actual use (44.40%), functionality (19.40%), reliability (27%), maintainability (20%), portability (40%), and operational acceptance/operation readiness (24%). The participants confirmed moderate satisfaction with efficiency (5.6%). Figure 7 shows the satisfaction indices as plotted, while Fig. 8 shows the main criteria weights.

These weights were further confirmed by the W_{ik}s that were calculated by linear programming optimal solutions previously shown in Table 8.

Table 10 Partial satisfaction index measurement (Note. Partial satisfaction is the mean of all measured variable at 0.385 (38.5%))

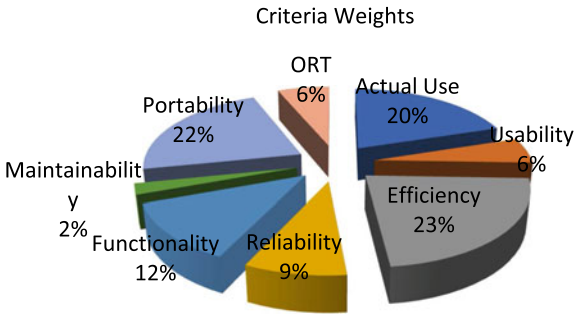


B2 f_x =1/100*((76/135*0)+(0/135*35.9)+(10/135*100)+(34/135*100)+(15/135*100))

	A	B	C	D	E	F	G	H	I
		$S_i = \frac{1}{100}\sum_{i=1}^{n} p_i^k x_i^{*k}$							
2	Actual Use	0.44							
3	Usability	0.81							
4	Functionality	0.19							
5	Efficiency	0.51							
6	Reliability	0.27							
7	Maintainability	0.20							
8	Portability	0.41							
9	ORT	0.25							
10									

Fig. 7 Satisfaction indices

Fig. 8 Main criteria weights



4.7 Demanding Index

The demanding indices were calculated following Eq. 5 as shown in Table 11.

A representation of the results obtained through the demand Eq. 5 is as shown in the chart presented in Fig. 9. In general, the higher the demand index, the more the satisfaction level that should be improved to fulfill the respondent's expectations. Moreover, Fig. 9 reveals that the sample consists of non-demanding users.

4.8 Impact Values

It is also important to know the ratio of improvement effort to the satisfaction level as satisfaction can result from a certain improvement effort. The improvement index is shown in Table 12. The improvement figures were converted to percentages and

Table 11 Demanding index measurement (Note. Demand index is the mean of all measured variables at 0.5447 (54.5%))

C7 =(SUM(((100*(1-1))/(5-1))-0)+(((100*(2-1))/(5-1))-C3)+(((100*(3-1))/(5-1))-C4)+(((100*(2-1))/(4-1))-C5)+(((100*(5-1))/(5-1))-C6))/SUM(100*SUM(0.25,0.5,0.75,1)))

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		Actual Use		Usability		Functionality		Efficiency		Reliability		Maintainability		Portability		ORT		
2		W[ik]	X[ik]	W[ik]	X[ik]	W[ik]	X[ik]	W[ik]	X[ik]	W[ik]	X[ik]	W[ik]	X[ik]	W[ik]	X[ik]	W[ik]	X[ik]	
3		2	2.8	35.9	0.9	26.3	2.5	32.5	2.7	20.0	0.7	17.3	0.1	18.4	0.0	0.0	0.4	13.2
4		3	5.0	100.0	2.5	100.0	5.2	100.0	10.9	100.0	3.5	100.0	0.6	100.0	7.6	100.0	2.7	100.0
5		4	6.8	100.0	2.0	100.0	4.2	100.0	8.5	100.0	4.1	100.0	0.7	100.0	6.8	100.0	2.9	100.0
6		5	5.4	100.0	0.7	100.0	0.2	100.0	0.9	100.0	0.7	100.0	0.6	100.0	7.6	100.0	0.0	100.0
7	Demanding	$D_i = \frac{\sum_{i=1}^{n}\left(\frac{100(k-1)}{q_i-1} - x_i^w\right)}{100\sum_{k=1}^{q_i-1}\frac{k-1}{q_i-1}}$	-0.51026		-0.603		-0.69		-0.51		-0.437		-0.573		-0.523		0.512	

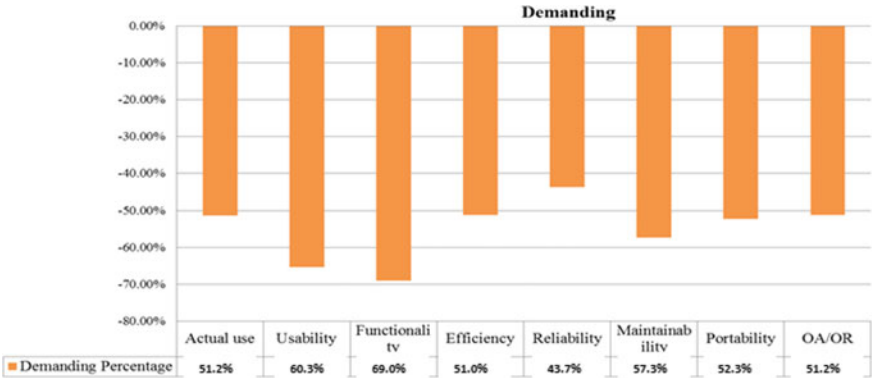


Fig. 9 Demand indices

represented in the impact chart as shown in Fig. 10, and the satisfaction indices shown in Fig. 7. The results show that a 40% increase in satisfaction can be achieved by improving actual use features by 11% while 2% of improvement efforts made to system maintainability, and 1% improvement on usability features will give a 20% and 80% increase in satisfaction respectively.

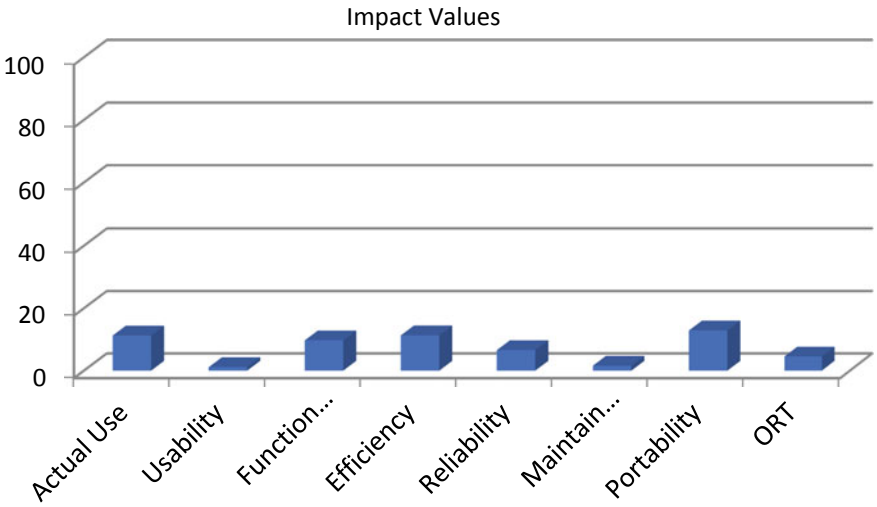


Fig. 10 Criteria impact values

Table 12 Improvement value function measurement

C2 =A2*(1-B2)

	A	B	C	D
1	Weight	Satisfaction Index	Improvement index $I=b_i(1-S_i)$	
2	0.20	0.4	0.11	
3	0.06	0.8	0.01	
4	0.12	0.2	0.10	
5	0.23	0.5	0.11	
6	0.09	0.3	0.07	
7	0.02	0.2	0.02	
8	0.22	0.4	0.13	
9	0.06	0.2	0.05	
10				

4.9 Improvement Values

Table 12 shows improvement value function measurement. For instance, using Table 12 and Fig. 10, improving usability features by 1% (serial #3 on the table) will result in an 80% level of satisfaction while improving functionality features by 10% (serial #4 on the table) yield 20% user satisfaction level (see Fig. 10). Furthermore, an 11% improvement of the efficiency features results in a 50% increase in satisfaction level, whereas a slight increase in reliability feature by 7% will yield a 30% satisfaction increase. Once more, a 40% increase in satisfaction level for system portability features is achieved by improving efforts by 13% while 5% improvement effort yields a 20% satisfaction increase in ORT. This means that users may not care much about these e-government systems meaning that they have not accepted them or they are ignorant of the value it adds to their lives. It is also an indication of non-demanding behavior.

5 Conclusion and Recommendation

User satisfaction is a critical quality measure of any product or service, especially software products and services [20, 31, 32]. It is also a key to user acceptance of products and/ or services. Moreover, for software systems and associated services, an evaluation based on rigorous models should enhance users' confidence and acceptance of the product and/or service. By applying a rigorous MUSA model in the measurement of users' satisfactions of any e-government system ensures that all critical concerns of the systems are taken into consideration in the evaluation process using a multicriteria approach. The outcome of this study with overall global satisfaction index of 86%, a partial satisfaction index of 54.5%, and a demand index of 38.5% suggests that the sampled users of Botswana e-government systems are generally satisfied with the systems, but there could still be minimal challenges especially with aspects that need improvement as indicated by the rating. Therefore, the MUSA

model gives interpretable results when applied approriately. The study among others makes the following recommendations:

- Public sensitization of the e-government systems and their expected benefits
- Government investment in human capital development through education and training of staff for proper management of the e-government systems
- Government should also explore alternative digital government development including mobile telephones, kiosks, and multi-channel access to services.
- Regular evaluation of the e-government systems is necessary as a way of ensuring that the systems deliver on their mandate.
- This study was conducted from the users' perspectives in government to citizens (G2C) systems. In the future, focus should include the perspectives of government to business (G2B) and government to government (G2G) with a higher number of participants (sample).

Acknowledgements The authors thank the Republic of Botswana, Ministry of Transport and Communication for granting permission for the study on which this paper is based [17].

References

1. Fenton NE, Pfleeger SL (1997) Software metrics: a rigorous and practical approach, 2nd edn. PWS Publishing, Boston, MA
2. Henderson-Sellers B (1996) Software metrics. Prentice-Hall, UK
3. Li W, Henry S (1993) Object oriented metrics that predict maintainability. J Syst Softw 23(1):111–122
4. Weyuker EJ (1988) Evaluating software complexity measures. IEEE Trans Softw Eng 14(1):1357–1365
5. Zuse H (1995) A framework for software measurement. Walter de Gruyer, New York
6. Schneidewind NF (2009) Systems and software engineering with applications. IEEE Press
7. Okike EU (2007) Measuring class cohesion in object-oriented systems using Chidamber and Kemerer metric and Java as case study. Ph.D Thesis, Department of Computer Science, University of Ibadan
8. Okike EU, Osofisan AO (2007) Validation of Chidamber and Kemerer's lcom metric using measurement theory. J Appl Inform Sci Technol 1(1):29–42
9. Avison D, Fitzgerald G (2006) Information systems development: methodologies, techniques and tools, 4th edn. McGraw-Hill
10. Oliver RL (1997) Satisfaction: a behavioral perspective on the consumer. Irwin/MacGraw-Hill, New York
11. Schuppa T (2009) E-government in developing countries: experiences from sub-Saharan Africa. Gov Inf Q 26(1):118–127
12. Rorrissa A, Demissie D (2009) The state of art of e-government services in Africa: an analysis of relevance. In: Proceedings of the 42nd annual hawaii international conference on system sciences, pp 1–8
13. Rorrissa A, Demissie D (2010) An analysis of African e-government services websites. Gov Inf Q 27(2):1–9
14. Sukhoo A, Kurrumchand S (2012) Towards effective e-government for Africa. In: IST-Africa Conference and Exhibition, pp 1–10

15. Dombeu JVF, Rannya N (2014) African e-government research landscape. African J Inform Syst 6(3):83–119
16. Viik L, Nyman-Metcalf K, Astok H, Viiderfeld T, Kaljurand K, Puua M (2018) The deployment of e-governance systems in Africa: final report. European Commission
17. Small ON, Okike EU, Tlhobogang B (2019) An evaluation of citizen satisfaction with e-government systems in Botswana: a multicriteria satisfaction analysis. In: Proceedings and book of abstracts of the 5th international conference on business innovation and growth, pp 83–103
18. Okike EU, Mosanako S (2020) Measuring customer satisfaction on software-based products and services: a requirements engineering perspective. In: Yang XS, Sherratt S, Dey N, Joshi A (eds) Fourth international congress on information and communication technology, vol 1027. Advances in Intelligent Systems and Computing. Springer, Singapore, pp 31–45
19. Lobadi NO (2021) A framework for evaluating user satisfaction of e-government systems in Botswana. Master of Science Information Systems Dissertation, Department of Computer Science, University of Botswana
20. Hill N, Brirley J, MacDougall R (2003) How to measure customer satisfaction, 2nd edn. Routledge, London
21. Sabri A, Loshini T (2014) Constructing a MUSA model to determine priority factors of a SERVEPERF model. Math Theory Modeling 4(11):33–43
22. Grigoroundis E, Siskos Y (2010) Customer satisfaction evaluation: methods for measuring and implementing service quality. Springer, New York
23. Grigoroundis E, Politis Y, Siskos Y (2002) Satisfaction, benchmarking and customer classification: an application to the branches of a banking organization. Int Trans Operations Res 9(5):599–618
24. Grigoroundis E, Litos C, Moustakis V, Politis Y, Tsironis T (2008) The assessment of user perceived web quality. an application of satisfaction bench marking approach. Satisfaction, benchmarking and customer classification: an application to the branches of a banking organization. European J Operations Res. Elsevier 187(3):1346–1357
25. Grigoroundis E, Arabatzis G (2010) Visitors' satisfaction, perception and gap analysis: the case of Dadia-lefkimi national park. Forest Policy Econ 12(3):163–172
26. Pallant J (2010) SPSS survival manual: a step by step to data analysis using the SPSS program, 4th edn. McGraw-Hill New York
27. ISO/IEC TR9126-4 Software engineering product quality—part4: quality in use metrics. International Organization for Standards, Geneva, Switzerland
28. Smith M, Thorpe R, Jackson P (2008) Management research, 3rd edn. Sage Publications, London
29. Statistics Botswana (2019) Quarterly multi-topic labour module. Quarter 1, Gaborone central statistics office
30. Heir JF, Black B, Babin J, Anderson RE (2010) Multivariate data analysis, 7th edn. Pearson, New York
31. Hagger MS, Chatzisarantis N (2008) Self determination theory and the psychology of exercise. Int Rev Sports Exercise Psychol 16(1):26–28
32. Jonathan D, Richard L (1992) A strategy for customer satisfaction. Cornell Hotel Restaurant Administration Q 35(5):32–40

Pain Detection Using Deep Learning Method from 3D Facial Expression and Movement of Motion

Kornprom Pikulkaew and Varin Chouvatut

Abstract Nowadays, face expression technology is widespread. For instance, 2D pain detection is utilized in hospitals; nevertheless, it has some disadvantages that should be considered. Our goal was to design a 3D pain detection system that anybody may use before coming to the hospital, supporting all orientations. We utilized a dataset from the University of Northern British Columbia (UNBC) as a training set in this study. Pain is classified as not hurting, becoming painful, and painful in our system. The system's effectiveness was established by comparing its results to those of a highly trained medical and two-dimensional pain identification. To conclude, our study has developed an uncomplicated, cost-effective, and easy to comprehend alternative tool for screening for pain before admission for the public in general and health provider.

Keywords Deep learning · 3D facial expression · 3D pain detection · Movement of motion

1 Introduction

Recent studies demonstrate that frequent monitoring of patient pain levels in hospitals may lead to considerable improvements in outcomes. This monitoring has been difficult to maintain due to the stress and strain already imposed on medical professionals, and an automated method may be a model therapy. Utilizing an automated facial expression system to do this is a possibility since pain may be specified using several facial action units (AUs) [1]. This investigation relied on data from the UNBC database [1], which includes sequences of photographs of people's faces as they do various range-of-motion assessments. FACS coders assigned an AU code to each frame in this data collection, containing self-report and observer assessments.

In the present, self-reporting techniques are often employed to measure pain. However, it has limitations between patients' and physicians' conceptualizations of

K. Pikulkaew · V. Chouvatut (✉)
Department of Computer Science and Graduate School, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand
e-mail: varin.ch@cmu.ac.th

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_67

pain. After that, an observer rating approach was developed to address difficulties with self-report measurements. The primary disadvantage of observer rating is that it is impractical and inefficient when monitoring is required for an extended time, such as watching a patient in a hospital. Additionally, the observer and self-report techniques of measurement are very subjective. Researchers [2] have suggested using facial expressions to identify pain as a possible solution to the challenges outlined above. Pain is quantified using the Facial Action Coding System (FACS) [3], accurately measuring pain. The data collected by such systems may be utilized to train a real-time autonomous system that provides significant benefits in terms of patient care and cost savings. The UNBC-McMaster Shoulder Pain Expression Archive is the most utilized resource for calculating pain detection algorithms.

The present research in facial pain detection focuses on improving accuracy via supervised and unsupervised learning [4]. Numerous studies have employed Support Vector Machine (SVM) [5] to categorize data to enhance the effectiveness of facial pain detection. However, none of the available approaches considered three-dimensional facial pain detection and motion tracking. Additionally, the purpose of our work is to assist medical personnel in spending less time monitoring individuals such as disabled individuals, toddlers, or ICU patients. Because young people and the disabled cannot talk to medical workers, our study aims to help lessen the stress of medical workers who have to keep an eye on everyone. Additionally, this procedure may save money since specific methods need the analysis of pricey tissues.

Our contribution will provide a unique approach for pain identification utilizing 3D facial expressions to enhance pain diagnosis using facial expressions. Second, 3D motion vectors that employ deep learning to classify data and exceed current state-of-the-art methods will be offered. The framework's results will be compared to the self-report at the end of the process. Additionally, doctors will be asked about their thoughts on the system.

2 Methodology

2.1 Pain Measurement Process

First, we used the UNBC dataset to train our system. The UNBC was developed by McMaster University and the University of Northern British Columbia. They took videos of the faces of people who had shoulder pain. A series of range-of-motion tests were performed simultaneously on both the afflicted and unaffected limbs, using active and passive methods.

Our 3D pain detection system quantified pain using AUs since they are dependable, accurate, and verified by physicians [6]. Prkachin discovered in 1992 that six action units contained the majority of information concerning pain. Six action units are included (4, 6, 7, 9, 10, 43). Additionally, we quantified pain using the Prkachin and Solomon pain intensity (PSPI) equation [7]. As shown in Fig. 1, our technique

Fig. 1 An illustration of a pain category derived from the PSPI

categorizes pain into three categories: no discomfort, becoming uncomfortable, and unbearable.

2.2 Process of Face Recognition Using a Deep Convolutional Neural Network

We emphasize accuracy above speed throughout the deep learning process. However, training time will be longer, prediction times will be longer, and memory requirements will be higher. ResNet-50 V2 is our infrastructure's face recognition model. He et al. [8] created the ResNet-50 V2 architecture. The number 50 denotes a layer in our 3D pain detection system's convolutional neural networks (CNNs). Additionally, the model uses transfer learning and pre-trained ImageNet weights [8].

There is always noise in the actual world. As a result, data augmentation is still required to optimize our system's performance. When photos are loaded for training purposes, our approach must crop the image to the middle square and scale it. Additionally, the system can automatically generate slight alterations in patient photos to account for the noise in real-world data. It created five picture variants with randomly changed brightness, contrast, saturation, hue, rotation, and scale during training. We trained on 80% of the sequenced photos and tested on the remaining 20%.

2.3 Stereoscopy

Sir Charles Wheatstone [9] pioneered stereoscopic imaging in 1838, and Sir David Brewster [9] later refined the technology. Stereoscopy, often known as stereoscopic imaging, is a method that creates the sense of depth in a flat picture. Stereopsis, often referred to as depth perception, is the visual perception of the relative distances between objects in one's field of vision. A few visual clues assist us in seeing objects in this manner. If one thing obscures another somewhat, we consider the object in front to be closer. As objects and patterns recede, and vertical lines converge, they become hazier and less saturated in color, shifting toward the blue end of the spectrum.

Stereoscopy is a technique for creating an illusion from a two-dimensional picture. The aim of using this strategy is to demonstrate the depth of the photographs. By creating a three-dimensional view that can be seen using a stereoscopic camera, 3D imaging is used to map aerial images and for amusement. The advantage of 3D is that it enables detailed observation and collects additional data to enhance the system's performance. Depth data is obtained from a two-dimensional picture by adding the depth map and merging it into the two-dimensional image, resulting in a three-dimensional image. The most straightforward stereoscopic system takes two 2D photographs of the same object while indenting the camera and shooting a little distance. Then, two different images were examined with each eye, adjusting the spacing to fit the image size. Our brain combines two pictures and converts them to a single vision with a shallow depth as if our two eyes were naturally seeing natural objects.

The advantages of employing two photographs side by side in this manner are that it simplifies production by requiring just two images to be put side by side. Additionally, it is viewable without the assistance of assistive technology. However, the disadvantages of this approach include the fact that the picture must be tiny, and the distance between the eyes must be set appropriately [10]. The devices used to view 3D photos are referred to as 3D viewers. According to technology, 3D viewers may be classified into active and passive. Dynamic models need an electrical system or electric current to switch on and off. In contrast, passive models are more superficial and lower in weight, relying just on modifying the filter in the eyeglass lens.

3 Result and Discussion

3.1 3D Pain Detection Using a Deep Learning Algorithm

This section utilized a novel 3D facial pain detection approach by converting two-dimensional coordinates to three-dimensional coordinates via stereoscopy. Thus, our method illustrates the world coordinate system in three degrees of freedom (translation in x, y, and z axes). To begin, we used 2D sequential images from the UNBC

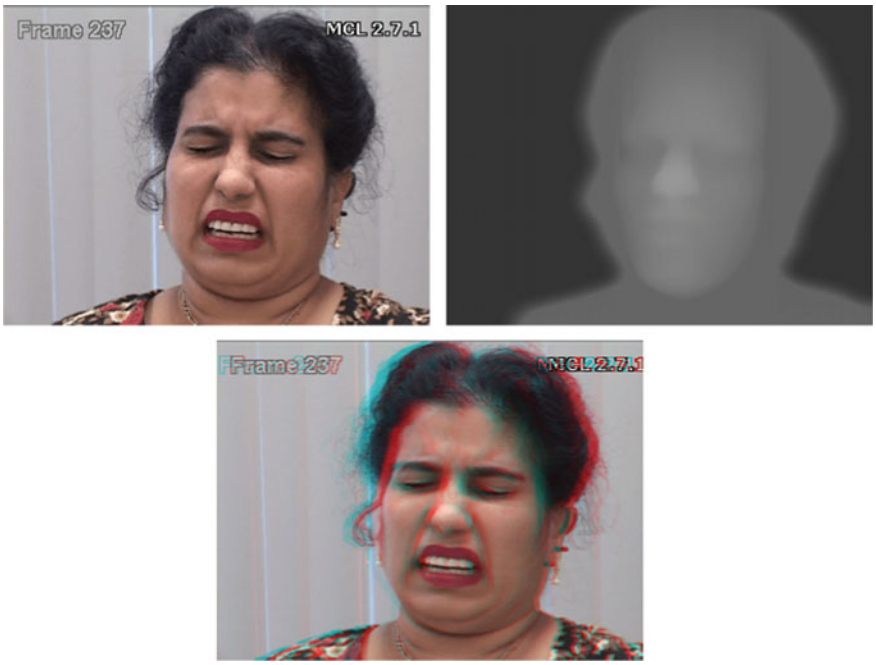


Fig. 2 An illustration of a pain category derived from the PSPI and corresponding stereoscopic content

dataset. Following that, we need to resize all the patient's images and adjust the lighting in the same way for each. For instance, we need to make each image 400 × 400 pixels and adjust the lighting in the same way. Second, we used the stereoscopy method to convert 2D images to 3D images. Finally, we obtained 3D images from which the 3D vector for defining each AU of facial expression could be calculated. Figure 2 summarizes the three-dimensional transformation process.

We used 3D technology because 3D facial pain detection has higher accuracy and performance than 2D facial pain detection. Table 1 illustrates the difference between 2D and 3D pain identification. Furthermore, as demonstrated in the author's previous work [11–13], 3D can support patients in any orientation that 2D cannot. The technique for detecting 3D pain is illustrated in Table 2.

3.2 Validation and Evaluation Procedures

Frequently, empirical data is the only method to determine if a model performs as predicted in trust and correctness. The accuracy, f1-score, and recall of a model are critical factors to examine throughout the assessment process. We verified our model

Table 1 Comparison of 3D and 2D pain technologies

3D	Accuracy (%)	2D	Accuracy (%)
(1) Initial pain	99	No pain	90.13 (misclassification)
(2) Initial pain	99	No pain	60.37 (misclassification)
(3) No pain	100	No pain	65.43
(4) No pain	100	No pain	93.68
(5) Pain	100	Pain	88.58
(6) No pain	99	No pain	92.74
(7) Pain	100	Initial pain	56.41 (misclassification)
(8) Initial pain	99	No pain	61.93 (misclassification)

using a variety of batch sizes and epochs in addition to classification criteria. The test results are summarized in Table 3.

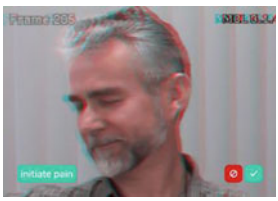
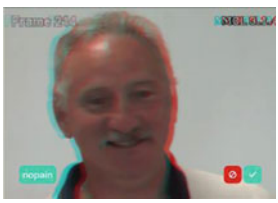
Therefore, the variables utilized for validation may be summarized as follows. Accuracy is expressed as the percentage of correct classifications. Precision refers to the rate of relevant outcomes. In contrast, recall refers to the proportion of relevant results that the model accurately categorizes.

4 Conclusions

We anticipate that 3D will eventually replace the two-dimensional method of pain identification due to its superior accuracy and performance. In the real world, we cannot urge people not to move while they are injured. Additionally, 2D cannot be employed since the accuracy will degrade and sometimes result in misclassification while the object is in motion. On the other hand, three-dimensional has a restriction that we must consider. For instance, 3D employed blue and red to suggest a complex image that the average person may not comprehend and would need the assistance of an expert to decipher.

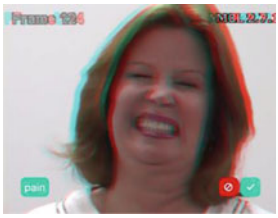
In future work, we intend to enhance the stereoscopy approach. It eliminates the requirement to convert 2D to 3D pictures by blue and red to indicate depth. The three-dimensional point cloud will be utilized to enhance our system. Regrettably, we cannot compare our approach to other academics since there is no 3D pain technology now.

Table 2 A demonstration of three-dimensional pain detection in human faces

Patients	Pain intensity	Confidences		
		IP	NP	P
	Initial pain	0.99	3.57×10^{-7}	6.50×10^{-9}
	Initial pain	0.99	1.03×10^{-7}	4.55×10^{-11}
	No pain	1.07×10^{-23}	1.0	1.87×10^{-31}
	No pain	4.86×10^{-15}	1.0	7.42×10^{-20}
	Pain	0.0	0.0	1.0

(continued)

Table 2 (continued)

Patients	Pain intensity	Confidences		
		IP	NP	P
	No pain	1.07×10^{-23}	1.0	1.87×10^{-31}
	Pain	1.78×10^{-29}	0.0	1.0
	Initial pain	0.99	1.20×10^{-6}	0.0009

IP = Initial pain, *NP* = No pain, *P* = Pain

Table 3 The result of our pain detection algorithm

Epoch	Batch size	Recall	F1-score	Precision (%)
10	32	0.97	0.97	97
	64	0.99	0.99	99
	128	0.98	0.98	98
15	32	0.99	0.99	99
	64	0.99	0.99	99
	128	0.98	0.98	98
50	32	0.99	0.99	99
	64	0.99	0.99	99
	128	0.99	0.99	99

Acknowledgements It was made possible, in part, by a grant from the Thailand Research Fund and Chiang Mai University under Thailand's Royal Golden Jubilee Ph.D. Program (grant number PHD/0226/2561). In addition, the author wishes to express his gratitude to the Computer Science Department and the Faculty of Science at Chiang Mai University for their unwavering support throughout the process.

References

1. Lucey P, Cohn JF et al (2012) Painful data: the UNBC-McMaster shoulder pain expression archive database. Image Vis Comput J 30:197–205
2. Hammal Z, Cohn JF (2012) Automatic detection of pain intensity. In: Multimodal interaction
3. Vijayanandh R, Balakrishnan G (2010) Human face detection using color spaces and region property measures. In: IEEE international conference on robotics and vision
4. Bargshady G, Zhou X, Wang H (2020) Enhanced deep learning algorithm development to detect pain intensity from facial expression images. Expert Syst Appl 149:1–10
5. Haque MA, Moeslund TB (2018) Deep multimodal pain recognition: a database and comparison of Spatio-Temporal visual modalities. In: Automatic face & gesture recognition
6. Hicks CL, Von Baeyer CL et al (2001) The faces pain scale-revised: toward a common metric in pediatric pain measurement. Pain 93:173–183
7. Prkachin KM, Solomon PE (2008) The structure, reliability and validity of pain expression: evidence from patients with shoulder pain. Pain 139:267–274
8. He K, Zhang X et al (2020) Deep residual learning for image recognition. In: Computer vision and pattern recognition
9. Wade NJ (2019) Ocular equivocation: the rivalry between Wheatstone and Brewster. Vision 3:1–11
10. Nityananda V, Read JC (2017) Stereopsis in animals: evolution, function and mechanisms. Experiment Biol 220:2502–2512
11. Pikulkaew K, Boonchieng E et al (2020) Pain detection using deep learning with evaluation system. In: Congress on information and communication technology
12. Pikulkaew K, Chouvatut V (2021) Enhanced pain detection and movement of motion with data augmentation based on deep learning. In: Conference on knowledge and smart technology
13. Pikulkaew K, Boonchieng W et al (2021) 2D facial expression and movement of motion for pain identification with deep learning methods. IEEE Access 9:109903–109914

The Next-Generation 6G: Trends, Applications, Technologies, Challenges, and Use Cases

Ayoub Bourbah, Bouchra Meliani, Zhour Madini, and Younes Zouine

Abstract While the 5G generation of mobile communications system has delimited its focus on Internet of Things (IoT) connection and industrial automation systems. The 6G generation will offer an extrasensory experiences through the fusion of the digital, physical, and human world. It will redefine the way we live, work, and manage the world, and it will make us more efficient thanks to the combination of intelligence and vigorous computation capabilities. The sixth generation of mobile communications is still under investigation. Several projects and research have been launched. 6G promises a specific type of communication with very high data rate and capacity, very low latency, maximum coverage, very high reliability, extremely massive connectivity, and very low cost and power. The current article provides an overview of what has been discussed about the future of 6G so far. The major goal of this paper is to present a comprehensive picture of 6G based on the research and projects that have been launched. We describe the potential architectural characteristics of 6G that will give users the experience they expect. We present an important list of technologies that will be the critical element in the rollout of 6G such as artificial intelligence, VLC communications, 3D beamforming, massive MIMO aircraft, and drones. We also exhibit scenarios and use cases that might be lived in this next-generation networks. Finally, we identify the challenges that could be faced by the 6G in different sides …

Keywords 5G · 6G · AI · MIMO · Data rate · Capacity · Massive connectivity · Maximum coverage · VLC · 3D · Beamforming

1 Introduction

Research related to 6G has already started; several entities have taken sides in this marathon, while the dream of 5G is still far from being global. This new generation of mobile telecommunication will be a revolution technology; it will bring new

A. Bourbah (✉) · B. Meliani · Z. Madini · Y. Zouine
Department of Electrical and Telecommunication, ISET Laboratory, Kenitra, Morocco
e-mail: bourbah.spe@gmail.com

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_68

type of communication (holographic, intelligent, and multidimensional) and exceed diverse fields of industry (telemedicine, robotics, marine communication …) [1]. In addition, the sixth generation will certainly deal with the shortcomings of the fifth generation (traffic issues, security, energy …). This paper represents a review article, which summarizes the various articles and white papers published under the 6G flag. In this article, in the first section, we will discuss the trends that spurred the emergence of the sixth generation. The second section presents the architecture of 6G characteristics and the possible technologies to ensure these characteristics. Sixth generation requirements and scenarios are described in Section III. At the end, we will present a variety of challenges facing the sixth generation.

2 Trends

The next generation of mobile communication (6G) receives starter cores from a variety of trends, we illustrate the most important:

2.1 Trafic

The world is moving to a fully connected and intelligent frame. The integration of diversified applications, such as artificial intelligence (AI), 3D services, Internet of Everything (IoE), and virtual reality, contributes to a massive quantity of information in whole the network [2]. In 2030, the traffic volume is estimated to reach 5016 EB/month [3], which it was just about 7.462 EB/month in 2010 [4] the massive production of data made the creation of a next generation of mobile communication system an essential goal [4].

2.2 AI

Finance, health care, manufacturing, industry, and wireless communication networks are just a few of the areas where AI is now being used [1]. The incorporation of artificial intelligence (AI) into wireless communication networks will increase the system's performance [1]. AI, for example, might: Enhance performance of handover operation [1] Improve network planning through location determination of the base station [1]. Reduce the consumption of energy—Permit self-healing of network anomalies (prediction and detection …) [1]. In the previous generations and specially the 5G generation, AI was not considered in the developing stage of the system. However, AI can be embedded in the initial phase of 6G conception [1].

3 6G Use Cases and Technical Requirements

6G will enable innovative, futuristic use cases that when deployed on a massive scale, will transform the way we live and work in remarkable ways.

3.1 Holographic Communications

Holographic displays, which provide three-dimensional images from one or many points to one or many points, are the coming step in the multimedia experience [5]. As a result, network interactive holographic capabilities will require a combination of extremely high data rates and extremely low latency [5]. The first is because a hologram is made up of numerous 3D pictures [5], whereas the second is because parallax is employed to allow the consumers to interact with the image as it varies according to his location [6].

3.2 Tactile and Haptic Internet Applications

This category contains various application examples:
Robotic and industrial automation The 6G arrival will usher in a manufacturing revolution, driven by networks that enable human–machine communication in cyber-physical systems (CPSs) [7], a vision known as Industry 4.0. This type of application necessitates communications between large connected systems without human intervention. As a result, robotics will need to be controlled in real time. A communication link with a maximum latency of 100s and a round-trip response time of 1 ms is required for advanced robotics scenarios in manufacturing [6].
Autonomous driving: (V2V or V2I) This sort of application has the potential to significantly reduce traffic accidents and congestion [8]. For collision detection and remote driving, however, delay on the range of a few milliseconds will most certainly be required [8]. The primary areas that 6G wants to enable are improved driver assistance, vehicle platooning, and completely autonomous driving [6].

3.3 Network and Computing Convergence

Similar to the 5G networks, edge computing architecture is expected to be used in 6G networks as well. When an user requests a service with low latency, the network directs the request to the closest edge computing point [5]. For computation-intensive applications and load balancing, a number of edge computing units could be engaged, but the computing resources must be deployed in a coordinated manner [5]. Edge

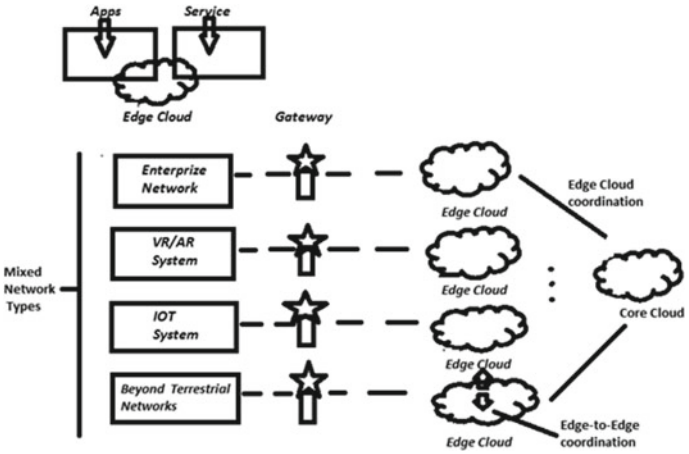


Fig. 1 Edge-to-edge coordination through local edge clouds [6]

cloud coordination is a good fit for AR, VR, autonomous driving, and holographic communications [6]. The major network requirements for this are computing awareness of constituent edge facilities, coupled network and computing resource scheduling (centralized or distributed), flexible addressing (any network node can become a resource provider), and rapid routing and rerouting (traffic should be able to route or reroute in response to load conditions) [6].

Figure 1 shows edge-to-edge coordination all over local edge clouds of various network and service categories, along with edge coordination with the core cloud architecture [5], to exemplify this concept [6].

3.4 Connectivity for Everything

We may discuss a variety of situations throughout this part, such as real-time monitoring of towns, the environment, transportation, and streets … [5]. As a result, the Internet of bio-things, enabled by smart wearable devices, and intra-body communications enabled by implanted sensors [5], will boost connection demand far beyond mMTC [9]. The most essential network needs for such situations are large aggregated data rates owing to vast amounts of sensory data, high security and privacy, especially when medical data are being transferred, and perhaps low latency when a speedy intervention (e.g., heart attack) is required. There are no mechanisms or models in place to evaluate these data requirements right now [6].

3.5 Chip-to-Chip Communications

Today, cable connections are used for on-chip, interchip, and interboard communication, so when data rates exceed 100–1000 Gb/s, those links become useless (bottlenecks). As a result, they must replace the cable connections with optical or THz wireless connections. Another attractive area for 6G is the development of such "nano-networks." Specific KPIs for nano-networks are dependent on chip implementations and applications [5] that will become apparent once they advance during the coming decade [6].

3.6 Space-Terrestrial-Integrated Networks

The concept of offering the Internet from space via enormous constellations of LEO satellites has gained popularity in recent years the connection to satellites instead of depending on existing terrestrial infrastructure become possible for mobile devices in this scenario [6].

4 Technologies

To ensure the good deployment of the network, 6G will be based on different and enhanced technologies:

4.1 Artificial Intelligence: AI

Artificial intelligence is one of the most essential technologies in 6G. In order to construct intelligent networks and real-time communication, this new generation will focus on machine learning technologies so that real-time data transit will be easier with the introduction of this technology, which will also boost efficiency and decrease data processing delay. For example, time-consuming steps (handover, network selection phase, etc.) can be executed quickly using AI [10].

4.2 Terahertz Communications

Because the RF band is currently not enough to fulfill the huge demands of 6G, this generation will be based on sub-THz communication and advanced mMIMO in order to increase bandwidth and spectral efficiency. The THz will play a key

part in 6G communication [4], which is intended as the next stage in high data rate transmission. The frequency band of THz communication is between 0.1 and 10 THz, which corresponds to the range 0.03–3 mm of wavelengths [4] but the frequency band recommended by ITU-R between 275 GHz and 3 THz, it is the main part for cellular communications [4]. As a result, by combining the THz and mmWave bands, the capacity of 6G cellular communication will be increased. At its smallest, the capacity is 11.11 times that of the mmWaves. We can get high data rates by employing the THz band [4]. THz band is define between 275 and 3 THz; it is composed from the mmWave 275–300 GHz and the far infrared 300 GHz to 3 THz. The important characteristics of using THz communication: The THz band can provide a large available bandwidth to deal with very high data rates. Due to the elevated frequency, the path loss will arise. So, in the 6G, the directional antennas will be needful because these antennas can reduce the interference. The devices and BS using the THz band will have a large number of antenna components [5], due to the wavelength of the THz waves [4].

4.3 Optical Wireless Technology

OWC technologies such as light fidelity, visible light communication (VLC), and FSO communication based on the optical band will be used in the future generation 6G [4]. Communication with optical wireless technology has a number of advantages, including increased data rates, reduced latencies, and provide secure communications [4].

4.4 FSO Backhaul Network

The data transmission in the FSO system is similar to that of optical fiber because the trans and receiver properties are identical to those of optical fiber [4], As a result, this technology is a good fit for delivering backhaul connectivity in 6G [5]. It is feasible to have quite long surface of communications with FSO, well over 10,000 km away. FSO provides high-capacity backhaul connectivity to far and non-remote locations including the all parts of the universe [5], as well as cellular BS connectivity [4]

4.5 Unmanned Aerial Vehicle

In 6G, BSs will be deployed in unmanned aerial vehicles (UAVs) or drones to enable high data rate wireless communication [5]; this technology has many advantages, like easy deployment and strong degrees of freedom with controlled mobility [4]. For example, in the case of natural disasters, or in unstable environments, drones

can easily manage these situations [4]. It can also serve to satisfying the requirements of three use cases: eMBB, uRLLC, and mMTC [4]. Furthermore, UAVs will help to improve network connectivity, detect incidents and disasters, and transfer information to emergency services.

4.6 Holographic Beamforming

Beamforming or channelforming is a signal processing technique used in antenna and sensor arrays for the directional transmission or reception of signals [4]. This procedure has many advantages, but the main property is the high signal-to-noise ratio and the high network efficiency [5]. Holographic beamforming is a new technique of beamforming for 6G that uses software-defined antennas [4]. This type of technology has more advantage compared to beamforming; HBF is able to transmit and receive signals in a more efficient and flexible way in multi-antenna communication [4].

5 The 6G's Architecture

As the generation of the 6G network is supposed to be a pure revolution in the field of communication, the targeted architecture should be able to meet the challenge. The services offered will be both diverse (XR, holographic communication, 3D communication, telepresence, etc.) and provided to an extremely dense population in different places in the universe (earth, ocean, air, space) which will need a strong management and dissemination capacity. To cope with these cases, the architecture of the 6G network should be characterized as follows (Fig. 2).

5.1 Global Architecture

The interconnection in the network will completely change with the 6G generation. We will move from a two-dimensional network to a multidimensional one. The coverage will be a global issue that is supposed to cover all parts of the universe (land, ocean, air, space, the micro world, and outer space) [11]. The 6G ground mobile communication system and the satellite communication system will form a network that gathers different type of communication (terrestrial/satellite/marine wireless communication) [11]. The 6G network will cover all of the areas of human activity: remote areas with no human living, deep oceans, and even the stratosphere [11]. Through different technologies such as visible light communication (VLC), free space optics communication (FSO Comm), and unmanned aerial vehicles (AUVs), the feature of global architecture can be assured.

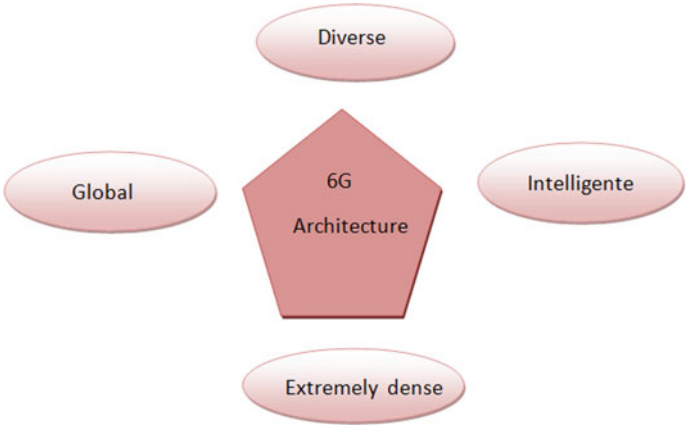


Fig. 2 The characteristics of 6G architecture

5.2 Diverse Architecture

The 6G was presented as an extra multi-service provider; the network should handle very diverse list of services at the same time for different users [12]. The services will be divided into large different categories as described in the use case section, depending on the capacity and the speed of the communication, the coverage extension, the energy consumption, and latency ….

5.3 Extremely Dense Architecture

As the number of connected devices is expected to reach 500 billion by 2030 [1], the 6G network architecture should be able to handle this high number of users; there will be different types of equipment (augmented reality (AR) glasses, virtual reality (VR) headsets, and holographic devices) with different quality of service criteria to be served at the same time. The 6G architecture must be prepared to serve this dense area of devices with sufficient throughput, acceptable speed, and an intelligent management unit [1].

5.4 Intelligent Architecture

Artificial intelligence (AI) will be included into 6G communication networks [5]. AI will be used to encompass all network sensors, control, physical layer signal processing, system administration, service-based communications, and so on [5].

The integration of AI during the development stage of communication systems will offer a super smart architecture, which will give more opportunities to take advantage of AI for improvement of overall network operation in terms of performance, cost, and ability to provide various services [1].

6 The Challenges Facing the Next Generation

As the 6G will bring a new vision in diverse sides (technologies, applications, and architecture …), this coming network will face multiple challenges in different levels.

6.1 The Terahertz Signal

THz frequency band will be the transmission features in 6G. However, it is difficult to work with such frequency range because: The generation of a continuous terahertz signals is tough due to the complexity of antenna/transmitter design. Thus, the generation cost will be high. The THz signals are characterized by a high energy loss. A Thz wave can be attenuated after a few meters. The presence of more moisture in the air make the loss energy much higher, consequently the 6G promises cannot be achieved [13]. The THz technology is still under investigations, and many researches are lunched to find an economical solution for these shortcomings.

6.2 The Energy

Energy is an important issue in the 6G. 6G will implement many technologies, which require high energy such as AI, edge computing, and big data. Consequently, it should find a solution for harvesting, charging, and conservation of energy. Signal processing mechanisms and data transmission required high power consumption too. In order to settle this energy issue, the 6G should offer new waveform and modulation mechanisms. The embedded AI in the 6G initial phase of development can efficiently reduce the quantity of energy consumed [13].

6.3 The Global Coverage

The global coverage goal, fixed by the 6G network, will be achieved through the low-earth orbit (LEO) satellite, having a height of 500–2000 km from the orbit. Its general purpose will be providing a global coverage with less path loss, lower trans-

mission delay. Nevertheless, the LEO satellite has many problems related to Doppler variation, Doppler shift, long transmission delay, and more path loss. As a result, many issues regarding synchronization, signal detection, and signal measurement may arise [13].

7 Conclusion

Given the expected increase in traffic by 2030, the 5G network will no longer be able to handle the big data transmitted. 6G will be the unique solution to deal with this problem and other expected challenges. Through this article, we have visualized the basic keys to 6G technology. We presented architecture and technologies, then requirements and scenarios, and concluded with the challenges that the sixth generation might deal with.

References

1. Samsung 6g Vision: research.samsung.com
2. Mathew A (2021) Chapter 10 sixth-gen wireless tech with optical wireless communication. Springer Science and Business Media LLC
3. Prakash V, Garg L, Azzopardi J, Camilleri T (2022) Chapter 5 High-speed connectivity. IGI Global
4. Chowdhury MZ, Shahjalal Md., Ahmed S, Jang YM (2020) 6G Wireless communication systems: applications, requirements, technologies, challenges, and research directions
5. arxiv.org
6. Tataria H, Shafi M, Molisch AF, Dohler M, Sjöland H, Tufvesson F (2021) 6G wireless systems: vision, requirements, challenges, insights, and opportunities
7. Li T, Khandaker MRA, Tariq F, Wong K, Khan RT (2019) Learning the wireless V2I channels using deep neural networks. In: IEEE 90th vehicular technology conference (VTC2019-Fall)
8. deepai.org
9. Tang F, Fadlullah ZM, Mao B, Kato N (2018) An intelligent traffic load prediction-based adaptive channel assignment algorithm in SDN-IoT: a deep learning approach. IEEE Internet Things J 5(6):5141–5154
10. DOCOMO White Paper 5G Evolution and 6G https://doi.org/10.1007/11823285_121
11. Lua Y, Zheng X (2020) 6G: a survey on technologies, scenarios, challenges, and the related issues
12. Zhao J (1907) A survey of reconfigurable intelligent surfaces: towards 6G wirelesscommunication networks with massive MIMO 2.0. arXiv:1907.04789
13. Nayak S, Patgiri R (2020) 6G communication: envisioning the KeyIssues and challenges. http://export.arxiv.org/pdf/2004.04024

Agility and Ambidexterity in SME—The Role of Digitization

Ralf Härting, Joerg Bueechl, and Jan Pach

Abstract The ongoing development and improvements of digital technologies lead to changes on markets and industries. The implementation of new digital approaches has the potential to secure the competitiveness and corporate success. Companies must anticipate these changes and respond to them adequately. As a result, organizations need to be more agile and ambidextrous. For further insights on these organizational requirements an empirical study has been conducted in conjunction with a literature review. The study examined whether digitization leads to an improvement in agility and ambidexterity, and thus to a better firm performance. This revealed drivers and inhibiting factors that have an influence on the potentials of agility and ambidexterity. Finally, a conceptual model has been developed to enable future quantitative-based research.

Keywords Digitization · Agility · Ambidexterity · Sensing · Responding · Exploration · Exploitation

1 Introduction

Advancements in digital technologies have led to uncertainty and volatility in different markets and industries [1]. The result is a more and more rapidly changing world, which is reflected by the widely known acronym VUCA. VUCA describes a world as volatile, uncertain, complex, and ambiguous [2]. To cope with the VUCA-world, agility has become increasingly important for companies to stay competitive and innovative [3]. Agility is essentially the combination of sensing and responding to business opportunities and unexpected changes in the market [4]. A lack of agility is

R. Härting (✉) · J. Bueechl · J. Pach
Aalen University of Applied Science, Beethovenstraße 1, 73430 Aalen, Germany
e-mail: ralf.haerting@hs-aalen.de

J. Bueechl
e-mail: joerg.bueechl@hs-aalen.de

J. Pach
e-mail: kmu@kmu-aalen.de

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_69

often the reason why companies fail to adapt to changing circumstances [5]. Companies, which survived in the present competitive environment, were often characterized by separating business units into explorative and exploitative ones. Such companies are defined as ambidextrous organizations [6]. Ambidexterity is frequently associated with improved organizational agility, respectively, performance. Furthermore, ambidextrous companies can better adapt to changing environmental conditions and in this context further research highlighted that digitization could potentially help to exploit the benefits arising from ambidexterity [5].

This is the basis of the investigation, which investigates the extent to which digitization promotes ambidexterity and agility in companies and what other driving and inhibiting factors exist in this context [6]. Not only the fact that SMEs seem to be more agile than large corporations [7] that they must deal with many limitations regarding resources, [8] but also the great importance of SMEs for the German economy justifies the study of German SMEs [9]. Due to the paucity of existing a qualitative research approach is used to subsequently conduct an inductive theory building.

2 Theoretical Framework

2.1 Ambidexterity

The term ambidexterity was first used by Duncan in 1976 [10]. Duncan focused on the ambidexterity in organizations, i.e., the organization is aligned toward today's business requirements, but remains adaptable to future changes in the environment [11]. Organizational ambidexterity enables two goals to pursue at the same time. On the one hand, to exhaust your own existing competencies or processes. On the other hand, ambidexterity can lead to new skills and new potentials can be discovered [12]. In 1991, March [13] described ambidexterity with the word's exploration and exploitation. According to Werder and Heckmann [10], exploration is about acquiring new knowledge to develop new products, open new markets and implement technologies. In contrast, exploitation is about improving existing knowledge and making small changes to create impact. Ambidexterity in form of exploration and exploitation does not only generate and promote innovations, competitive advantages, and prosperity, but mainly contribute to the company's success and thus to the survival of the organization [14]. In 1991, March [13] stated that exploration and exploitation are handled equally in a company. The problem was that both directions competed for the same resources, which created paradoxical tensions between exploration and exploitation [10]. These tensions must be resolved by compromising toward exploitation, which leads to short-term profits, but at the expense of the company's effectiveness and sustainability [6]. Tushman and O'Reilly state that this is the only way to ensure the long-term success of the organization [15] by simultaneous pursuit of increasing

one's own competencies and efficiency, as well as the simultaneous acquisition of new knowledge and research into new innovations [14].

2.2 Agility

The term and concept of agility dates back to the 1960s, when agility essentially described the adaptability of the manufacturing industry and its processes [16]. In the late 1980s and 1990s, agility emerged in the business research context and described organizations, which respond instantly to changing consumer and market expectations with product and process innovation [17]. In the early 2000s, the concept of agility came more and more into focus with the emergence of agile project management. Well-known examples are SCRUM, KANBAN, and Design Sprints [18].

Since then, a variety of different approaches regarding agility have been developed. Despite the divergent approaches, most regard agility as the capability of an organization to sense and respond to environmental changes with a relatively high speed and to take advantage of new opportunities [20, 27]. Three main characteristics can be identified. The first one is the sensing ability, which enables companies to detect environmental change. The second one is the ability to seize opportunities that become available due to those sensed changes. The final one is the responding ability, which helps organizations to respond more rapidly and effective as well as efficient [21, 22].

Tallon et al. [22] mention four categories of enablers of agility. The first category deals with technology. With IT resources and IT infrastructure, companies can react faster, more specifically and more easily. The second category focuses on management, which must pay attention to sensing, responding and dictate action. In the third category, agility is promoted by the corporate structure and organization itself with the focus on corporate strategy, business models and decision-making paths. The fourth category deals with the environment in which the company is located [22].

2.3 The Link Between Digitization, Ambidexterity, and Agility

Companies must sense changes and opportunities in the market or in their environment and adapt and align their resources and skills accordingly [19]. With a look at the ambidexterity in this context, it is noticeable that especially the exploration fits the characteristics of sensing and responding [6]. Because exploration means recognizing these market opportunities and acquiring new skills. This is done through a reorientation of the previous skills and competencies [14].

Through digitization, an organization can recognize environmental changes and trends more quickly and react to them more flexibly. From this context, the terms IT sensing and IT responding emerged [1]. Ambidexterity also benefits from advancing

digitization. On the one hand, through IT exploitation, existing technologies are used and improved to increase efficiency and effectiveness. On the other hand, new technologies and competencies are developed during IT exploration. Thus, through digitization and the subsequent ambidexterity, companies can optimize their costs and ultimately improve their firm performance [1].

The present work deals with the analysis of SMEs with focus on digitization, agility, and ambidexterity. This analysis was chosen because there is a research gap in the literature [4]. The following research question should be answered by using qualitative data:

> How does digitization in SMEs influence agility and ambidexterity and what driving and inhibiting factors can be identified?

3 Methodology

3.1 Data Collection

The data collection for this research project includes ten semi-structured interviews with ten different companies and ten different industries in 2020 and 2021. Participants from different functional areas and levels were interviewed to gain a broad understanding of the situation and to minimize individual bias. 30% of respondents hold management positions. Table 1 lists the partners of each interviewee. Semi-structured interviews are deliberately used because they can maintain a high degree of consistency between questions and offer a good comparability between interviews

Table 1 Overview if interviewees

Firm	Interviewee	Gender	Department/position	Industry	Employees
1	E-Commerce 1	Male	Managing Director	E-Commerce	50
2	Service 1	Male	Head of Humas Resources	Breakdown Service	280
3	Data 1	Male	Project leader R&D	Internet a Multimedia	72
4	Production 1	Female	Marketing	Sauna Manufacturer	734
5	Consulting 1	Male	Service Team Leader	Consulting	70
6	Finance 1	Female	Project Management Organisation	Banks and Finances	992
7	Textile 1	Female	Human Resources	Textile and Fashion	880
8	Service 2	Male	Head of Marketing	Crane Service	170
9	Pharma 1	Male	Chief Information Officer	Pharma	950
10	Wholesale 1	Female	Assistant Purchasing Manager	Wholesale	880

Source Own representation

without preventing the discovery of unknown and unexpected phenomena [24]. The interview lasted an average of 45 min and was recorded digitally, and then transcribed verbatim in the original German language using the software Amberscript [25].

3.2 Data Analysis

The data have been analyzed with the Grounded Theory approach. Therefore, first open coding, then axial coding and lastly selective coding was applied. In the open coding phase basic common codes are assigned. These regulations can be derived from direct statements in interviews or from constant comparisons with the literature [26].

For example, the codes "Accessibility of Data and Service" and "Digital Tools" have been merged into first order category "Digital Transformation", and then combined with the first order categories "Employees" and "Leadership" into the higher order category "Drivers of agile and ambidextrous working methods". To retrace the method, there is an excerpt of the codes and categories provides in Fig. 1. During the analysis, theoretical memos were written, which helped to cycle back and forth between literature and collected data to ensure validity. Transcribed data were analyzed by using the software for qualitative research MAXQDA.

4 Empirical Findings and Model Conceptualization

Based on a previous study [27] and the analysis of the empirical qualitative data we collected on agility, ambidexterity, and digitization in German SMEs, we developed a conceptual model, which is depicted in Fig. 2.

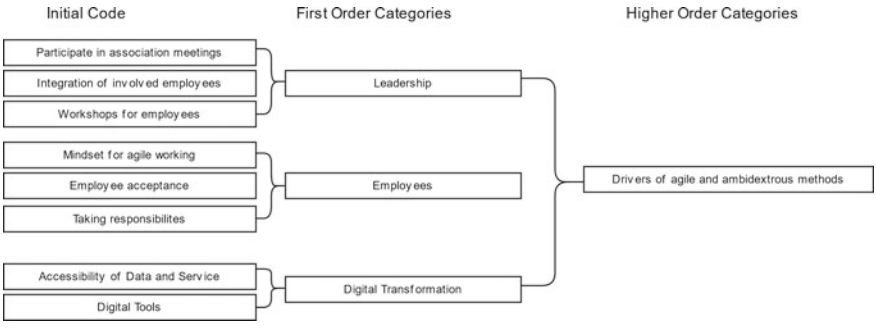


Fig. 1 Exemplary procedure of the data analysis. *Source* Own representation

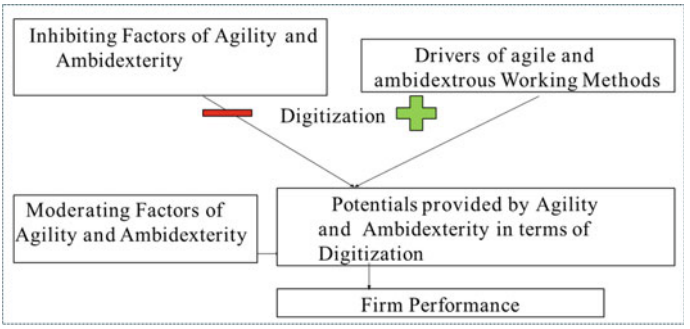


Fig. 2 Conceptual model of potentials of agility and ambidexterity

4.1 Drivers of Agility and Ambidexterity

Leaders can increase agility with increasing employee acceptance of change and with improving their understanding of the need for change. Besides Leaders, employees are another important driver for agility and ambidexterity. It is especially important to involve employees in projects from the very beginning to give them the feeling that they can help shape the project. The acceptance of change from the perspective of employee plays an important role. Furthermore, an agile mindset of employees must be considered as another driving factor. In this context, employees have to be encouraged to become active where it is not clear whether it will work. At this point, Pharma 1 mentioned the mindset of a culture of constructive criticism leading to a learning culture. Therefore, the willingness to work simultaneously on different aspects must be seen as an important prerequisite for working ambidextrously in SMEs. Furthermore, digital transformation enables employees to access the data they need from anywhere at any time and improve the accessibility of Data and Service, as Production 1 emphasized.

4.2 Inhibiting Factors of Agility and Ambidexterity

Employees can be both drivers and inhibitors of agility and ambidexterity. Many employees do not understand why they should do anything differently after many years, as Service 2 noted. Especially the knowledge of employees can inhibit an agile mindset, because in some cases the knowledge limits the solution space and agile development opportunities remain unrecognized, as Pharma 1 emphasized. Another inhibiting factor can be seen in the misunderstanding between employees. When it comes to ambidexterity in SMEs, the available resources often prove to be an inhibiting factor. Above all, the need for a simultaneous approach emerged in almost every interview conducted. In addition, (incorrect) development and implementation

of digital solutions can become an inhibiting factor for agility and ambidexterity. As e-commerce 1 pointed out, any process that is to be implemented digitally consistently and according to a fixed pattern leads to a certain rigidity and less agility. Furthermore, according to Consulting 1, it does not make sense to simply digitize existing processes; instead, the opportunity should be taken to rethink or optimize existing processes and workflows. Another problem is that employees do not use digital solutions optimally or use them incorrectly, which prevents them from working quickly and efficiently. In addition, insufficient implementation of digitization projects is another inhibiting factor.

4.3 Potentials for Agility and Ambidexterity Provided by Digitization

For agility, the major potentials must be seen in designing processes more flexible and to enable faster work. Digitization enables companies to respond more flexibly to changes. The Corona pandemic was frequently mentioned in correlation with this. As Service 1 noted in this context, digital tools make it possible to work independently of location and time. Besides flexibility, speed is the most important reason why companies strive for digitization and agility. In addition to faster communication, faster finishing of tasks and faster availability of required data were mentioned. Regarding the faster availability of data, Pharma 1 mentioned the advantage of being faster in troubleshooting, since, for example, historical data can be accessed more quickly, and possible parallels can be unveiled. In this context, faster research of information plays an important role in saving time and reducing the workload of the whole team.

4.4 Moderating Factors

In this context, the branch structure of the companies turns out to be a restraining instrument of agility and ambidexterity. The pharmaceutical industry does not see itself as hostile to innovation, but as very conservative and risk averse. Furthermore, some sectors are not per se designed for agile working. A further aspect with regards to departments as a moderating factor is the continuous digitization with its Data Privacy and its Digital Technologies. Companies are more connected than ever before with cloud applications and digitized workflows. IT systems are open to the outside world. This offers many advantages, but also makes them vulnerable to attack. The other factor which moderates the influence of the potentials on firm performance, is the increased pace of changes. Pushed by the Corona pandemic, many companies have offered home offices to their employees. This in turn drives ambidexterity, as an exploratory approach focuses on innovative and creative approaches.

5 Discussion

The potentials identified in this study are an increase in flexibility, a faster way of working and a higher efficiency in the way of working. In addition, the branch structure, the continuous digitization as well as the increasing speed of change turn out to be moderating factors. If this conceptual model will be considered in the implementation of digitization projects, agility and ambidexterity in SMEs can be increased and better business performance achieved. The long-term benefits of digitization are time savings, competitiveness, and efficiency. The study has also certain limitations. Future research should quantitatively verify or disapprove the established hypotheses mentioned in this paper. In addition, different homogeneous groups should be interviewed to obtain a higher validity and representativeness. In general cross-sectional studies would benefit from a long-term research design. The terms agility and ambidexterity are largely unknown in SMEs beyond the IT business function. Therefore, agile and ambidextrous approaches, which have been carried out rather tacitly and partly unconsciously in SMEs, might be identified with alternative research designs.

References

1. Leonhardt D, Haffke I, Kranz J, Benlian A (2017) Reinventing the IT function: the role of IT agility and IT ambidexterity in supporting digital business transformation
2. Bennett N, Lemoine G (2014) What a difference a word makes: understanding threats to performance in a VUCA world. Bus Horiz 57(3):311–317
3. Teece D, Peteraf M, Leih S (2016) Dynamic capabilities and organizational agility: risk, uncertainty, and strategy in the innovation economy. Calif Manage Rev 58(4):13–35
4. Chan CML, Teoh SY, Yeow A, Pan G (2018) Agility in responding disruptive digital innovation: case study of SME. Inform Syst J 29(2):436–455
5. Rialti R, Marzi G, Silic M, Ciappei C (2018) Ambidextrous organization and agility in big data era: the role of business process management systems. Bus Manage Process J 25(5):1091–1109
6. O'Reilly CA III, Tushman ML (2008) Ambidexterity as a dynamic capability: resolving the innovator's dilemma. Res Organ Behav 28:185–206
7. Dibrell C, Davis PS, Craig JB (2008) Fueling innovation through information technology in SMEs. J Small Bus Manage 46(2):203–218
8. Li L, Su F, Zhang W, Mao J (2018) Digital transformation by SME entrepreneurs: a capability perspective. Inf Syst J 28(6):1129–1157
9. Audretsch DB, Lehmann EE, Schenkenhofer J (2018) Internat. Strategies of hidden champions: lessons from Germany. Multinatl Bus Rev 26(1):2–24
10. Werder K, Heckmann CS (2019) Ambidexterity in information systems research: overview of conceptualizations, antecedents, and outcomes. J Inf Technol 20(2):28–52
11. Gibson C, Birkinshaw J (2004) The antecedents, consequences, and mediating role of organizational ambidexterity. Acad Manage J 47(2):209–226
12. Hsu C-W, Lien Y-C, Chen H (2013) International ambidexterity and firm performance in small emerging economies. J World Bus 48:58–67
13. March J (1991) Exploration and exploitation in organizational learning. Organ Sci 2(1):71–87
14. Goel S, Jones III R (2016) Entrepreneurial exploration and exploitation in family business: a systematic rev. and future directions. Family Bus Rev 29:94–120

15. Raisch S, Birkinshaw J (2008) Organizational ambidexterity: antecedents, outcomes, and moderators. J Manag 34(3):375–409
16. Harraf A, Wanasika I, Tate K, Talbott K (2015) Organizational agility. J Appl Bus Res 31(2):675–686
17. Gergs HJ (2019) Agilität und Organisationsentwicklung – Ziemlich beste Freunde? Wie die Organisationsentwicklung die digitale Transformation von Unternehmen unterstützen kann und wie sich dabei selbst verändern muss. In: Gruppe. Interaktion. Organisation. Zeitschrift für angewandte Organisationspsychologie (GIO), vol 50(2)
18. Kaim R, Härting RC, Reichstein C (2019) Benefits of agile project management in an environment of increasing complexity—a transaction cost analysis. In: Czarnowski, Ireneusz/Howlett, Robert J./Jain, Lakhmi C. (Hrsg.) Proceedings of the 11th KES international conference on intelligent decision technologies (KES-IDT 2019), vol 2. Springer, pp 195–204. ISBN 978-981-13-8303-8 (eBook), ISSN 2190-3026 (electronic)
19. Weber Y, Tarba SY (2014) Strategic agility: a state of the art, introduction to the special section on strategic agility. Calif Manage Rev 56(3):5–12
20. Overby E, Bharadwaj A, Sambamurthy V (2006) Enterprise agility and the enabling role of information tech. Eur J Inform Syst 15(2):120–131
21. Guangya S (2011) Exploring requirements of agility for knowledge management. In: 6th conference on professional knowledge management—knowledge to action, pp 371–381
22. Tallon PP, Queiroz M, Coltman T, Sharma R (2019) Information technology and the search for organizational agility: a systematic review with future research possibilities. J Strat Inf Syst 28:218–237
23. Tallon P, Pinsonneault A (2011) Competing perspectives on the link between strategic information technology alignment and organizational agility: insights from mediation model. MIS Q 35(2):463–486
24. Myers MD (2019) Qualitative research in business and management
25. Locke K (2001) Grounded theory in management research
26. Glaser BG, Strauss AL, Strutzel E (1968) The discovery of Grounded Theory, strategies for qualitative research. Nurs Res 17(4):364
27. Bueechl J, Härting RC, Pressl M, Kaim R, Reichstein C Potentials and barriers of agility in small and medium sized enterprises. In: Abramowicz W, Sören A (eds) 24th international conference on business information systems, TIB Open Publishing Business Information Systems, vol 1, pp 367–380

Maximizing the Score and Minimizing the Response Time in Scrabble Game

Alok Singh Gahlot, Vashista Bhati, and Ruchi Vyas

Abstract The perfection of information is an important notion in games when considering sequential and simultaneous game. It is a key concept when analyzing the possibility of strategies and decision making for the next move. Scrabble board game is an imperfect information game where each opponent is unaware of the letters on the opposite rack. It is board game of luck, vocabulary and strategy. Our computer board game strategy focuses on maximizing the score and minimizing response time while searching for the best word possible from letters through the dictionary of words where best word could be a word containing all letters of rack and with best scoring letters (Vyas and Gahlot in A theoretical approach to reinforcement learning algorithm and Nash equilibrium to maximize the score in imperfect information game. Springer, Singapore, 2021) [1].

Keywords Decision making · Imperfect information game · Scrabble

1 Introduction

In games research there has always been a focus on strategy being used by opposite player and predicting their game plan but there are several areas in games which are yet to be explored on improving the score and minimizing the searching time for the best move [3]. There have already been researches done on perfect information game but there is still a huge scope in imperfect information game just like scrabble board game. Optimizing the time spent while finding correct word with good score is the main goal of this research.

A. S. Gahlot (✉) · R. Vyas
MBM College, Jodhpur, India
e-mail: aloksg@gmail.com

V. Bhati
Department of Technical Education, Government of Rajasthan, Jodhpur, India

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_70

1.1 Scrabble Overview

Scrabble board game has square boxes on a board which is a 15*15 matrix. A bag containing 100 different letters is used in game from those 10,098 tiles have alphabet and their respective points, ranging from 0 to 10, 2 blank tiles which are wild tiles which could become any alphabet as per requirement by player. When a blank is played, it will stay in the game as the letter it fills in for various letters in the game will have different points. Game can have 2 or 4 players playing in teams or individually. Players draw 7 letters from the bag and arrange them on their racks and make words when they get chance [2]. After each move or word, they put on board they get score according to the alphabet and the position of board they have kept the letters as board has several positions like double letter, double word, triple letter, triple word score. If a player makes word from all 7 letters of the rack it is said to be bingo with 50 bonus points. Scrabble is one of the brilliant board games that remain unique to the complete society [4]. There are several researches being done on retention of player in scrabble game to make it more interesting so that no player leaves in between the game [6] (Fig. 1).

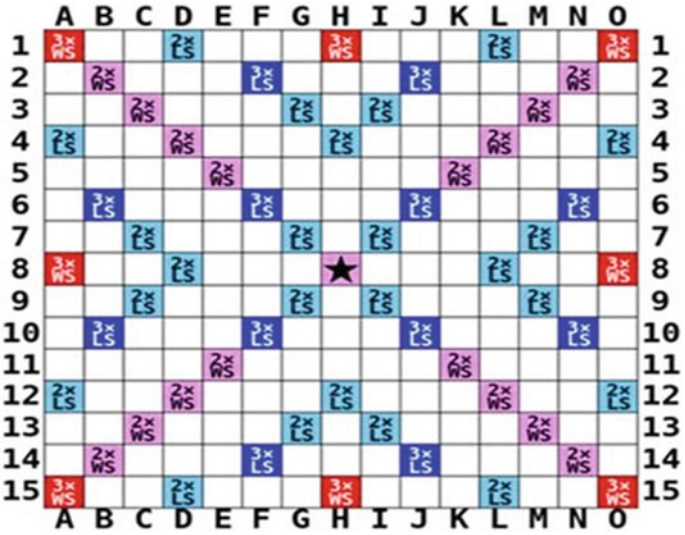


Fig. 1 A scrabble board game. The red-colored tiles specify triple word score, dark blue specifies triple letter score, turquoise specifies double letter score, and pink specifies a double word [1]

2 Related Work

There are several research being done on imperfect information games like Scrabble till date, and some of them have already been recognized and are being used by gaming platforms.

Opponent Modeling by Richards et al. [3] this exploration works by foreseeing the adversary, in light of Bayes' hypothesis, penances exactness for effortlessness and simplicity of calculation. However, even with this worked on model, we show critical improvement in play over a current Scrabble program. These exact outcomes recommend that this straightforward estimate might fill in as a reasonable substitute for the immovable somewhat recognizable Markov decision. This work centers around computer versus computer Scrabble play, the devices created can be of extraordinary use in preparing people to play against different people.

The five-point per game benefit against a non-inferencing specialist is likewise huge from a down to earth point of view. To give the distinction some specific circumstance, we performed correlations between a couple of sets of systems. The benchmark technique is the calculation which is consistently pick the move that scores the most points on the current turn. An expert that joins a static leave assessment into the positioning of each move overcomes an eager player by a normal of 47 points per game. At the point when a similar Static Player goes up against Quackle's Strong Player, the mimicking specialist wins by a normal of around 30 points per game. To have the option to average five points more for each game against such a tip top player is a serious significant improvement. In a competition setting, where standings rely upon wins and misfortunes as well as spot on spread, five extra points per game could have a huge effect. The improvement acquired by adding adversary displaying to the recreations would appear to legitimize the extra computational expense.

Scrabblesque by Harrison and Roberts [2] works for Player Retention in Scrabble, this examination chips away at how to drive player after specific stages in game player loses its advantage so this exploration utilizes AI game adaptivity and game investigation to build maintenance time of player up to 11.3%.

Man-made reasoning (AI) has a long-standing and sound connection with games where it turns into a famous application region for AI-driven exploration like game playing, game plan, etc. With the new headway of AI game playing programs that had surpassed human abilities, decency turns into a significant issue to be addressed to guarantee the appeal of a game can be held to the future players. Such an issue may likewise be intensified while considering the setting of turn-based games, where the principal player might enjoy a gigantic upper hand over the ensuing player(s) (called the upside of drive). It proposes a creative method for making a game appealing while at the same time keeping up with reasonableness by taking on the Komi (remuneration framework). This decency arrangement is approved by applying the framework onto a word re-arranged word game, Scrabble, where reasonableness can be kept up with dependent on the expertise level of the players [4].

AI did not perform well for final stages, and the AI lost to Quackle normally. Quackle AI could simply outscore Maven moderate final stage AI [7]. Finally, No

Q-Sticking AI was found to be a promising AI during mid games. Nevertheless, this AI did not perform well during final stage circumstances. The critical disadvantage of this research was the execution time taken by the three-handle look-ahead.

3 Method

Implementation of this game is done by initializing a 15*15 Matrix which has DL, DW, TL, TW at respective positions and start position. Initially it asks for number of players who want to play the game 2/4, then it asks for mode you want to play in easy or hard as in easy mode our game plays normal game while for hard mode our game uses proficient tricks to win the game. when user chooses the mode scrabble_move() function works which checks the rack and get their rack filled with 7 letters from bag of tile automatically, now player 1 gets the first chance and it plays for any word made from his rack, now player 2 which is our player calls a function get_next_move() and this function loads the dictionary and calls another function throne() which returns alphabet from the board which have space around them so that a word could be made now that word with letters of the rack creates permutation of words, now these words are checked against the scrabble dictionary for valid words now after receiving valid words from dictionary all the words are passed on to another function score_find() which gives off board value of the word now comparison is being done on score while calling score_word() function which returns the highest scoring word with its onboard value. Now our player plays the maximum scoring word while calling the function throne. Each time any player plays for a word it gets the board position through the function square_shape() as well as the time spent while finding the word and placing on board and the score of both players. Finally a function call is made on board_shape() which gives the winner of the game and average response time of both players. The strategy used here is to reduce the response time of searching the word and increasing the score.

Our computer checks the permutation of words against a text file which contains the highest scoring before checking to the main scrabble dictionary as these are the words with 7–8 letters as it wants to score for bingo majority of the time this the strategy where we get word from the 30 words with highest score in scrabble time to search for valid word would be reduced and it will also be the highest scoring word as well. This game has created this text file while playing the game several time as whenever it played the game and made a bingo it stores that word in a text file named my_dict.txt which is created after certain simulations and we have received around 30 words in text file.

This strategy maximizes the score as majority of the time our player is trying to find bingo score and searching time will be reduced as we are finding in small set of words hence this implementation will enhance the score as well reduce the response time which is found to be 1.8 s on an average and score of around 264 on an average as well as on a scale of 10 games it wins 60–70% of the times it plays (Fig. 2).

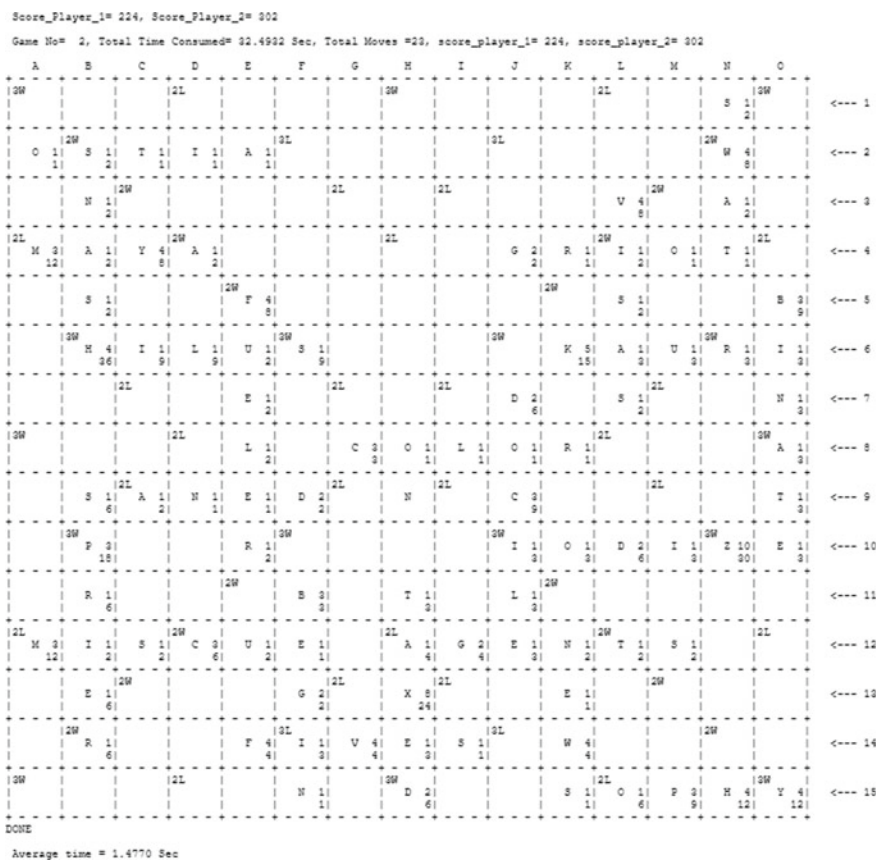
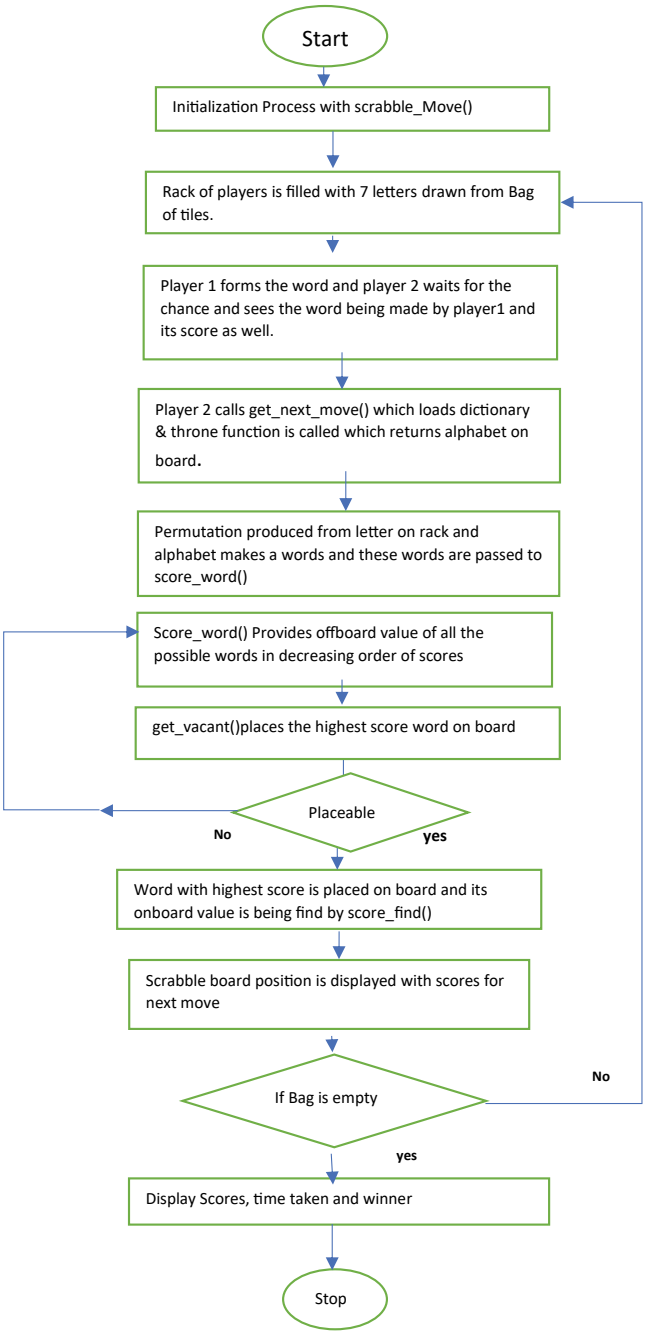


Fig. 2 Screenshot of the output of 1 of the game

Flowchart


Start
Initialization Process with scrabble_Move()
Rack of players is filled with 7 letters drawn from Bag of tiles.
Player 1 forms the word and player 2 waits for the chance and sees the word being made by player1 and its score as well.
Player 2 calls get_next_move() which loads dictionary & throne function is called which returns alphabet on board.
Permutation produced from letter on rack and alphabet makes a words and these words are passed to score_word()
Score_word() Provides offboard value of all the possible words in decreasing order of scores
get_vacant()places the highest score word on board
Placeable
No
yes
Word with highest score is placed on board and its onboard value is being find by score_find()
Scrabble board position is displayed with scores for next move
If Bag is empty
No
yes
Display Scores, time taken and winner
Stop

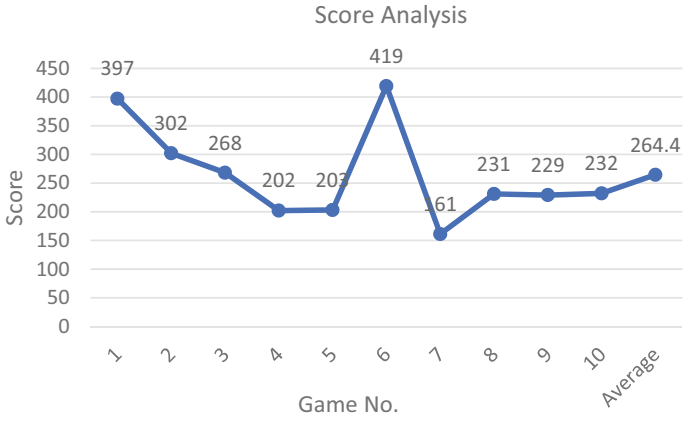


Fig. 3 Graph showing different scores achieved by our player

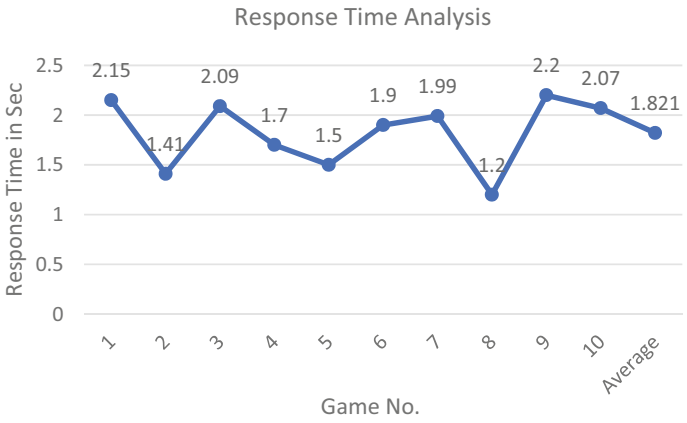


Fig. 4 Graph showing response time of our player

4 Experiment Results

See Figs. 3 and 4; Tables 1 and 2.

5 Conclusion

The game implemented here is majorly focusing on winning ratio and therefore it is trying to increase the score by creating a text file which has 30 words which uses 7/8 letters which is supposed be creating bingo on scrabble board which makes 50 points

Table 1 Showing the result of 10 games played

Games	Response time (s)	Score	Win
1	2.15	397	Win
2	1.41	302	Win
3	2.09	268	Win
4	1.7	202	Lose
5	1.5	203	Lose
6	1.9	419	Win
7	1.99	161	Lose
8	1.2	231	Win
9	2.2	229	Win
10	2.07	232	Win
Average	1.821	264.4	70%

Table 2 Comparison of various on winning percentage

Game	Win %
Maven	60
Quackle	40
Scrabble[a]	70

[a]Implementation of scrabble used in this research

so our game initially checks that whether our permutation has a bingo word from the text file and if it is there it makes that word which reduces the time of searching from main dictionary and increases the score too.

This implementation has enhanced the score, reduced the response time and winning percentage. Conclusion to this approach is that when our player against any human it may score around 264 as an average score while making a word from the 7 letters on the board could take around 1.4 s while on a scale of 10 games played it wins 7 times.

6 Future Work

There is further more scope of increasing the score and minimizing the response time, if machine learning algorithm such as q learning or Monte Carlo tree search to implement the scrabble game which is being used widely now a days in game playing while using Nash equilibrium for decision making could enhance the statistics as the game will be improvised if you use different strategy to make it a better version.

References

1. Vyas R, Gahlot AS (2021) A theoretical approach to reinforcement learning algorithm and Nash equilibrium to maximize the score in imperfect information game. Springer, Singapore.https://doi.org/10.1007/978-981-16-3915-9_6
2. Harrison, Roberts DL (2013) Analytics-driven dynamic game adaption for player retention in Scrabble. In: 2013 IEEE conference on computational intelligence in games (CIG), pp 1–8.https://doi.org/10.1109/CIG.2013.6633632
3. Richards M, Amir E (2007) Opponent modeling in scrabble. In: IJCAI international joint conference on artificial intelligence, pp 1482–1487
4. Aung HPP, Khalid MNA, Iida H (2019) Towards a fairness solution of Scrabble with Komi system. In: 2019 international conference on advanced information technologies (ICAIT), pp 66–71.https://doi.org/10.1109/AITC.2019.8921355
5. Shapiro SC, Smith HR (1988) A scrabble crossword game-playing program computer games I, pp 403–419
6. Aiolli F, Palazzi CE (2019) Enhancing artificial intelligence in games by learning the opponent's playing style. In: IFIP international federation for information processing, pp 1–10 [online]
7. Cheng C, Zhu Z, Xin B, Chen C (2017) A multi-agent reinforcement learning algorithm based on Stackelberg game. In: Conference: 2017 6th data driven control and learning systems (DDCLS), p 727. https://doi.org/10.1109/DDCLS.2017.8068163
8. Vrabie D, Lewis FL (2011) Integral reinforcement learning for finding online the feedback Nash equilibrium of nonzero-sum differential games. Adv Reinf Learn. https://doi.org/10.5772/13804
9. Abass AAA, Mandayam NB, Gajic Z (2021) Evolutionary random-access game with objective and subjective players. IEEE Access 9:35562–35572. https://doi.org/10.1109/ACCESS.2021.3062212
10. Whitehouse D, Cowling PI, Powley EJ, Rollason J (2013) Integrating Monte Carlo tree search with knowledge-based methods to create engaging play in a commercial mobile game. In: Proceedings artificial intelligence and interactive digital entertainment conference, Boston, Massachusetts, pp 100–106
11. Peng J, Williams RJ (1994) Incremental multi-step Q-learning. In: Machine learning. Morgan Kaufmann, pp 226–232
12. Dahl FA (2001) A reinforcement learning algorithm applied to simplified two-player Texas Hold'em poker. In: Proceedings of the 12th European conference on machine learning. Springer
13. Miikkulainen R, Bryant B, Cornelius R, Karpov I, Stanley K, Yong CH (2006) Computational intelligence in games. IEEE Computational Intelligence Society, Piscataway, NJ
14. 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, 64–67

Survey on Precision Agriculture in Indian Context for Effective Fertigation Using Learning Techniques

Bhagwan Dinkar Thorat and Sunita A. Jahirabadkar

Abstract The biggest issues found in fostering economic potential and yield value of precision agriculture is intelligence because of this we are not able to apply innovative techniques to complete variety of task online and offline. As agriculture affects long chain of supply management from crop production till its delivery to end user, it can be seen as field witch is going to produce high number of jobs. As India's economy is hugely depend upon the field of agriculture there is need to includes technology due to which it should be able to increase its 15.40% share of total crop production at the global level. Therefore, the need for an hour is to optimize the output per unit drop of water. Therefore, in today's sense, great emphasis is put on enhancing irrigation practices in order to increase crop production and preserve productivity levels. Use of advanced fertilizer irrigation system is going to strengthen root of the current agricultural system. The reduction in water usage due to the drip irrigation system varies from 30 to 70% for the surface irrigation method and the productivity benefit ranges from 20 to 80% for the various crops. Fertigation is a process of delivering water soluble fertilizer at the time of irrigation. Use of the fertigation process is going to significantly reduce the work load of the farmers and will save huge amount of time. Fertigation enables the placement of NPK nutrients directly into the plant root zone at the necessary dose during critical times. It is possible to increase crop yield capacity by three times more with the same amount of water by incorporating drip fertigation. Application all the above techniques will help in the process of quality and the production enhancement of the yield.

Keywords Fertigation · NPK · Crop yield · Learning technologies

B. Dinkar Thorat (✉)
Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon, Pune, India
e-mail: bhagwan.thorat@gmail.com

S. A. Jahirabadkar
Department of Computer Engineering, Cummins College of Engineering for Women, Pune, India

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4_71

1 Introduction

The world's population is growing day by day. The growing population will also increase the demand of food in the coming days [1]. Due to increasing urbanization, farm land is currently declining. Therefore, modern farming is required for those who need to get maximum crop yield in minimum farm lands. In country like India agriculture plays most important role in economy sector [2]. In India, 75% of rural people and their families depend on agriculture. Indian agriculture contributes about 15–16% of the total GDP. India is an agricultural country and its crop yield is less than other countries and the government of India is implementing various schemes to address these issues. Given India's current agriculture and economic situation and government policies, it is imperative to use modern farming or smart farming. It will help farmers to increase their crop yields and minimize the cost of inputs and losses due to reasons such as uncertain weather and climate conditions that have often affected crop production, agriculture and animal breeding [3]. To survive the changing conditions of the Indian economy, the agricultural sector needs a major upgrade. The proper combination of existing techniques with technological advancement are going to boost farming [4]. By using remote sensors, satellite images, surveys, etc., this information is obtained. In order to exploit me this data should be readily available to farmers along with the expertise of subject experts and researchers. There is also a dire need to analyse it to extract useful facts and patterns as the amount of such information is gradually increasing. This is where technology and computer science come into the picture. For this reason, many algorithms have been proposed over time, resulting in better results.

Twenty-first century gave us the concept of precision farming. Precision farming is a technique of using scientific advanced approaches in the field of agriculture. Farmer using GPS can be seen as example of precision farming. Precision farming is all about during the farming process at the right place on right time. This is highly going to affect the correctness of the precision farming is intended to match agricultural practices according to agro-climatic conditions.

Fertigation is a method in which fertilizer is dissolved, diluted and spread using a drip irrigation system along with water [5]. It primarily uses liquid fertilizers and water soluble fertilizers. It improves the productivity of fertilizer usage from 80 to 90% (Figs. 1 and 2).

$$\text{Precision Agriculture} = \text{Data collection} + \text{Data analysis} + \text{Implementation}$$

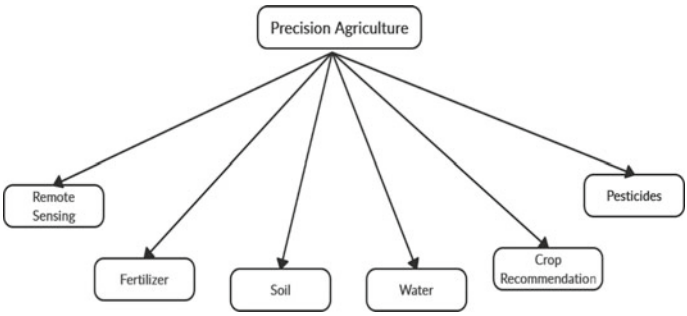


Fig. 1 Precision agriculture parameters

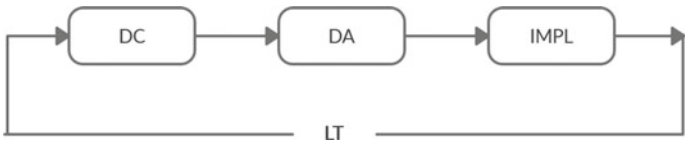


Fig. 2 Precision agriculture

2 Motivation

1. To bridge the gap between the current and desired capabilities of agricultural by developing an adaptable learning techniques to implement for precision agriculture.
2. To increase food production using fewer resources and reducing production cost.
3. To reduce socio-economic problems such as environment pollution, farmers suicides.

3 Literature Survey

Lakhiar et al. [1] reported that in agriculture, smart sensor techniques are very helpful. In agriculture, it is used to correctly organize certain acts and the process with the help of minimum amount of resources and farmer intervention. Several steps during plant cultivation with factors like water supply, moister in the air, temperature with several other factors will help plant growth significantly. Aeroponics is emerged as a highly popular method across the globe. It enable farmer to track the various affecting parameters without use of laboratory and also enabled farmer to navigate the whole system remotely.

Mohamed et al. [6] reported soil humidity content is very essential factor from the perspective of water management in precision farming. There are several factors

with should we taken into account whilst considering the soil moister contents. The research done at Nile's delta to understand prediction of various affecting factor from the perspective of total yield.

Doshi et al. [4] reported that Indian farmers select the crop they are going to take highly on their intuitions can be consider as huge mistake. The belief is due to the old farming techniques witch are passed on generation to generation which makes farmer very help less from the perspective on the yield as they have to highly relies on the uncontrollable factors like weather and soil conditions. All the above factors are very important and should be taken into the account by a single farmer. The farmer's single misguided or imprudent decision could have undesirable repercussions for both himself and the region's agricultural economy.

Pudumalar et al. [7] the common mistake made by Indian farmers is ignoring the soil needs whilst choosing a crop with is highly going to affect total yield. They are facing a significant productivity setback because of this. This farmers' issue has been solved by precision farming. This decreases the incorrect option of a crop and increases productivity.

Hey, Zhang et al. [8] the author should that crop classification and identification is very essential phase at the beginning witch is going to affect the whole process of farming. Various classification techniques can be used to make a whole structural and systematic overhauling.

Zheng et al. [9] claimed that in order to increase the total yield, the natural factors of specific location should be duplicated in order to obtain locally imported goods. It uses CropDeep datasets to provide data for the construction of a classification and detection model based on deep learning based on practical agricultural task characteristics in the greenhouse. Data sets are obtained by different greenhouse cameras and equipment that are more useful than previous datasets in providing management services for precision agriculture.

Yalcin [10] noted the phonological monitoring of agricultural plants. Accurate identification of phonological plant changes that would have enhanced the timing of the harvest, control of pests, yield prediction, monitoring of farms, disaster warning, etc. Usage of Deep Learning Architecture to identify and classify phonological phases of many plant types solely on the basis of visual data collected by cameras every half an hour. CNN helps us to understand the various factors of images.

Philipe et al. [11] stated peak period for harvesting of apple trees. Bloom intensity should be studied theory as it is most important factor as the production of apple. The CNN should be properly gauged. Carlos et al. [12] stated the method for the protection of crop from the hazardous chemical use. He also proposed new technique based on ML which is going to give us number of items with less computation efforts and correctness.

Erler et al. [13] stated for a spatially resolved LIBS spectrometer used determination of nutrients and different parameters of the soil in two areas of agriculture and main nutrients Ca, K, Mg, N, P. The determination of many major and minor nutrients was the focus of this work. It was possible to determine many nutrients with high accuracy.

Elasbah et al. [5] According to the source, nitrogen fertilizers used in irrigation are likely to raise nitrate levels in groundwater. The HYDRUS-2D/3D model help us to understand the movement of fertilizer. During the simulation various types of fertilizer were taken into the account. The effect of various precision farming techniques are also examined.

Leena and Saju [14] stated various identifiers and classifier to classify macronutrient deficiencies in maize plants, as described. Two identifiers techniques were implemented on the same leaf for two different feature sets and algorithms are implemented on the maize plant to determine micronutrients need.

Summary of Literature Survey

Author	Subject	Summary	Technology	Limitations
Lakhiar [1]	Monitoring and control system in agriculture	The aeroponic system is a modern agriculture's new plant cultivation process	Intelligent sensor techniques aoreponic	Crop rotation is going to give guaranteed growth in the yield
Mohamed [6]	Calculating humidity and its relationship crop	Predict humidity count and crop pattern with ORS and SAR	SMC, LST and NDVI triangle method	Need to improve more data set with large no of attributes To explore to improve the more accuracy by using other techniques like machine learning or deep learning classification and detection frameworks
Doshi [4]	Intelligent crop recommendation system using machine learning	The choice of which crop to grow is influenced by a number of environmental and geographic factors	Big Data Analytics	Crop rotation is going to give guaranteed growth in the yield
Pudumalar [7]	Crop recommendation system	Farmers will be assisted in sowing the appropriate seed based on soil requirements in order to maximize production and benefit from such a technique	Random tree, CHAID, K-nearest	Need to improve data acquisition techniques

(continued)

(continued)

Author	Subject	Summary	Technology	Limitations
Zhang [8]	Crop fine classification	Focusing on crop classification and identification is an very essential pre requirement from the perspective of production productivity	Hyperspectral remote sensing	The technique of hyperspectral crop is going to help to improve in the process of classification and identification of various affecting factors
Zheng [9]	Classification and detection in precision agriculture	Induction of novel technique in greenhouse	(CNN) YOLOv3	Over all important of classification process and detection tools from perspective of agriculture application
Yalcin [10]	Plant phenology	Recognize and identify phonological stages of various plant species solely based on visual data collected every half-hour by cameras	Convolution neural networks	Create own CNN architecture that is specifically trained to distinguish plant phonological levels
Diasa [11]	Apple flower detection	A novel method for detecting apple blossoms accurately	CNN	To extend the applicability from the perspective of flowers
Carlos [12]	Prediction of plant protection product deposition	With low computational cost and high precision, predicts the volume of pulverized materials	ML	To extend investigation with help of various available technologies
Erler [13]	Soil nutrient detection	LIBS spectrometer used for a spatially resolved determination of essential factors in relation to soil nutrients Ca, K, Mg, N, P	Regression methods	Improvement will require huge amount of data to performed without a significant loss of information in the multivariate regression

(continued)

(continued)

Author	Subject	Summary	Technology	Limitations
Elasbah [5]	Modelling of fertilizer transport	Understand the life cycle of fertilizer with the found techniques	HYDRUS-2-D/3-D model	Investigation of different learning techniques for increase efficiency
Leena and Saju [14]	Classification of macronutrient deficiencies in maize plant	Find macronutrient deficiencies	ANN, SVM, KNN	It can be used to predict various factors causing various disease

4 Gap Analysis

During the literature survey, it is observed that there is a lack of studies of the effectiveness of precision agriculture in Indian context. It is also observed that very few studies had applied the idea of the drip fertilizer system using different learning technology. Many of these focuses on research studies in other developed countries. As India has diverse environmental conditions, it is much challenging task to perform research in Indian context. Due to which there is a scarcity of the dataset. The literature survey also gives the different choices of modelling techniques available in order to investigate this phenomena.

5 Proposed Approach

Via drip systems, the proposed system utilizes drip fertilizer. Drip irrigation enables fertilizer application to be versatile, as it is simple to add fertilizers through the irrigation water. The roots in the wetted area have been found to improve their water and nutrient absorption ability. Fertigation is a process of delivering water soluble fertilizer at the time irrigation. Liquid fertilizer is used in this form, as well as water soluble fertilizer. This approach improves the productivity of fertilizer usage from 85 to 95%. Fertigation process helps to reach the required nutrients to the roots. Fertigation has the ability to ensure that the correct combination of water and nutrients is available in the root region, meeting the total water and nutrient requirements of the plant. The secret to high yield is the correct combination of water and nutrients.

Three variables require commitment to effective fertigation scheduling.

1. Specific nutrient needs of the crop and location.
2. Timing the distribution of nutrients to fulfil crop needs.
3. To reduce leaching of soluble nutrients below the successful root region, monitor irrigation (Fig. 3).

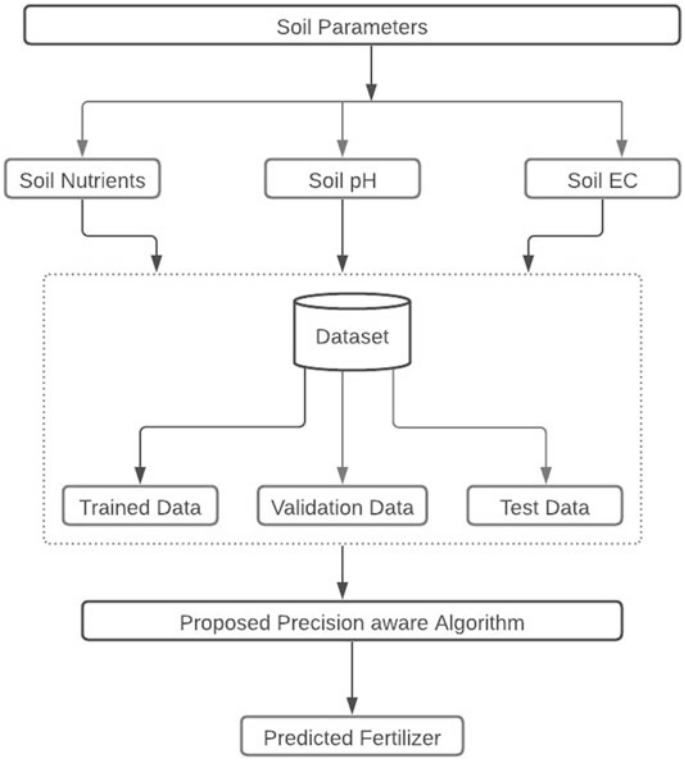


Fig. 3 Proposed architecture

6 Issues and Challenges

Indian farming also faces a number of issues, some of which are natural, whilst others are man-made. Any of these topics are debated.

The division of the agricultural region into small fragments is solely the responsibility of our traditional inheritance laws. The land belonging to the father is divided fairly amongst his sons, and so this process goes on.

In this way, with each passing generation, the assets are smaller and more dispersed. The fragmentation of the holdings of land would damage the farm's productivity. The main problems resulting from the fragmentation of the land are irrigation problems.

Agriculture is India's very ancient art, and Indian soils have been used for crop cultivation for thousands of years without taking much note of its replenishment. This has resulted in soil exhaustion and degradation, resulting in low productivity.

For almost all crops, the average yields are the lowest in the world. The need for fertilizers and manure for the land poses this dilemma. For good yields, nourished

soil is essential. However, in promoting the use of fertilizers in Indian agriculture, there are many practical challenges.

We all know that Indian agriculture, which is unpredictable, unstable and volatile, is dependent on the monsoon. The need for a proper irrigation system emerges from this problem.

Despite development in agriculture, the majority of farmers use traditional ploughing, sowing, irrigation, thinning and harvesting tools. The majority of human labour is used by marginal and small farmers, which results in the wastage of human labour and low per capita labour yields.

In order to enable farmers to implement technically advanced agricultural machinery, the government is making agriculture in India prosperous.

Soil erosion is one of Indian agriculture's most troubling issues. It is Indian agriculture's biggest evil. The soil is the most important commodity, ensuring good crop yields through its productivity. Yet soil erosion causes an immense loss of soil productivity.

7 Conclusion

India is an agriculture dominated country. If farmers are going to succeed automatically India will prosper. Our work is for achieving the above by inducing the technological advancement to improve the traditional farming process. The farmer can therefore use the right fertilizer to cultivate using soil parameters to increase its yield and thus increase the nation's overall productivity. The challenges and promising solutions listed above are predictions for the future of Indian agriculture.

References

1. Lakhiar IA et al (2018) Monitoring and control system in agriculture using intelligent sensor techniques: a review of the aeroponic system. Hindawi J Sens
2. Suchithra MS, Pai ML (2019) Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters. Elsevier
3. Bhanumathi S, Vineeth M, Rohit N (2019) Crop yield prediction and efficient use of fertilizers. IEEE
4. Doshi Z et al (2018) AgroConsultant: intelligent crop recommendation system using machine learning algorithms. IEEE
5. Elasbah R et al (2019) Modeling of fertilizer transport for various fertigation scenarios under drip irrigation. MDPI
6. Mohamed ES et al (2018) Mapping soil moisture and their correlation with crop pattern using remotely sensed data in arid region. Elsevier
7. Pudumalar S et al (2016) Crop recommendation system for precision agriculture. IEEE
8. Zhang Y et al (2019) Advances in crop fine classification based on hyperspectral remote sensing. IEEE
9. Zheng Y-Y et al (2019) CropDeep: the crop vision dataset for deep-z learning-based classification and detection in precision agriculture. Sensor

10. Yalcin H (2017) Plant phenology recognition using deep learning: deep-pheno. IEEE
11. Philipe A, Tabb A, Medeiros H (2018) Apple flower detection using deep convolution networks. Elsevier
12. Krishnan A et al (2017) Plant lea disease detection using deep learning and convolutional neural network. IJESC
13. Yu H et al (2010) A neural network ensemble method for precision fertilization modelling. Elsevier
14. Leena N, Saju KK (2018) Classification of macronutrient deficiencies in maize plant using machine learning. IJECE

Author Index

X.-S. Yang et al. (eds.), *Proceedings of Seventh International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 464,
https://doi.org/10.1007/978-981-19-2394-4