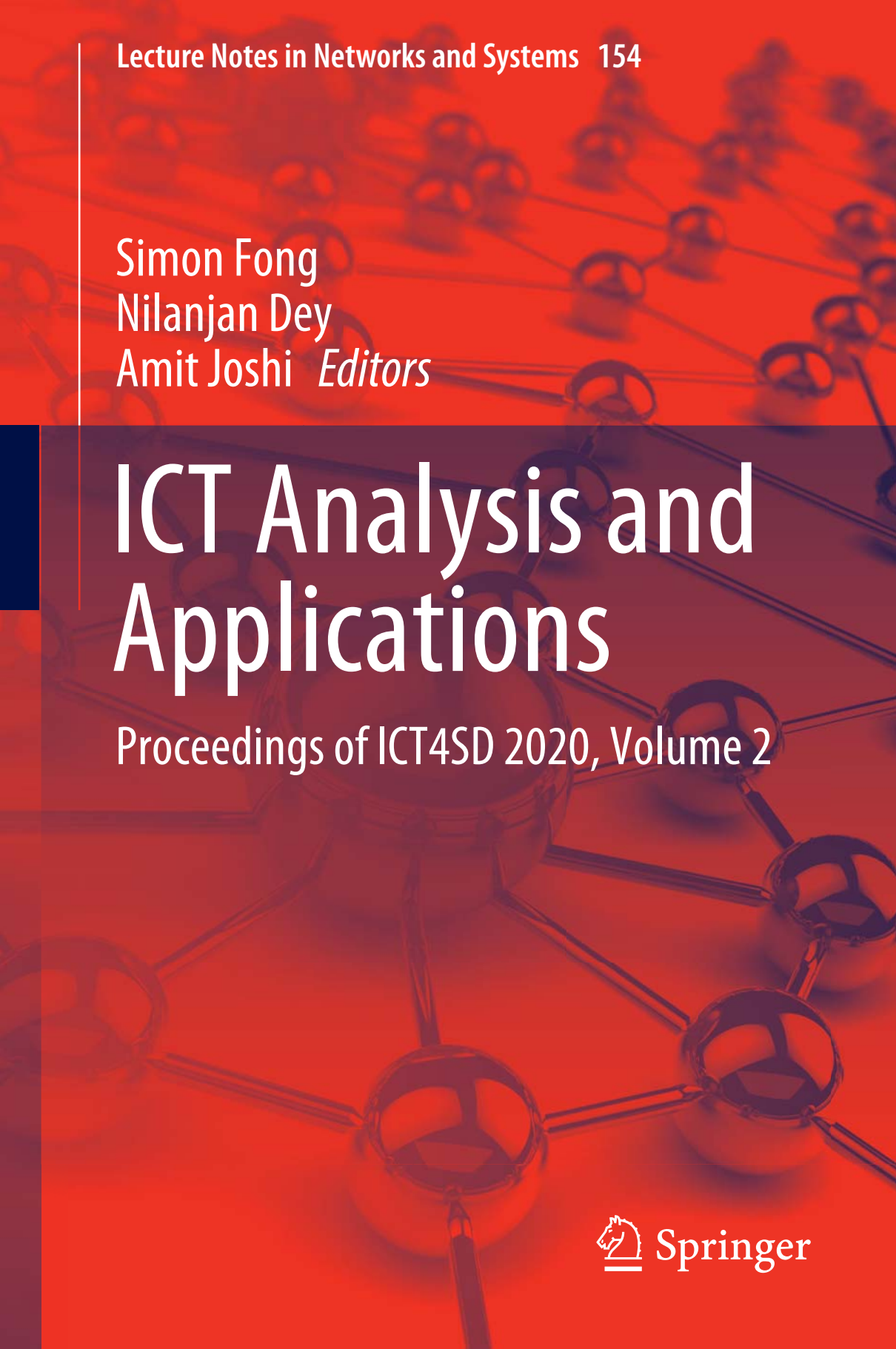
Lecture Notes in Networks and Systems 154

Simon Fong
Nilanjan Dey
Amit Joshi *Editors*

ICT Analysis and Applications

Proceedings of ICT4SD 2020, Volume 2



Springer

Lecture Notes in Networks and Systems

Volume 154

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.

More information about this series at http://www.springer.com/series/15179

Simon Fong · Nilanjan Dey · Amit Joshi
Editors

ICT Analysis and Applications

Proceedings of ICT4SD 2020, Volume 2

Springer

Editors
Simon Fong
University of Macau
Zhuhai, Macao

Nilanjan Dey
Techno India College of Engineering
Kolkata, India

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-15-8353-7 ISBN 978-981-15-8354-4 (eBook)
https://doi.org/10.1007/978-981-15-8354-4

Preface

Fourth International Conference on ICT for Sustainable Development (ICT4SD 2020) targets theory, development, applications, experiences and evaluation of interaction sciences with fellow students, researchers and practitioners.

Conference may concern any topic within the conference scope. Workshops may be related to any topics within the conference scope. The conference is devoted to increase the understanding role of technology issues, how engineering has day by day evolved to prepare human friendly technology. The conference will provide a platform for bringing forth significant research and literature across the field of ICT for sustainable development and provide an overview of the technologies awaiting unveiling. This interaction will be the focal point for leading experts to share their insights, provide guidance and address participant's questions and concerns.

The conference was to be held during 23–24 July 2020 at Hotel Vivanta by Taj, Panaji, Goa, India, but due to the pandemic this year, it was organized through digital mode. The conference was organized by Global Knowledge Research Foundation, Supporting Partner InterYIT, International Federation for Information Processing, State Chamber Partner Goa Chamber of Commerce & Industry, and National Chamber Partner Knowledge Chamber of Commerce & Industry.

Research submissions in various advanced technology areas were received after a rigorous peer review process with the help of programme committee members and 187 external reviewers for 1000+ papers from 19 different countries including Algeria, USA, United Arab Emirates, Serbia, Qatar, Mauritius, Egypt, Saudi Arabia, Ethiopia, Oman out of which 160 were accepted with an acceptance ratio of 0.15.

Technology is the driving force of progress in this era of globalization. Information and communication technology (ICT) has become a functional requirement for the socioeconomic growth and sustained development of any country. The influence of information communications technology (ICT) in shaping the process of globalization, particularly in productivity, commercial and financial spheres, is widely recognized. The ICT sector is undergoing a revolution that has momentous implications for the current and future social and economic situation of all the countries in the world. ICT plays a pivotal role in empowering people for

self-efficacy and how it can facilitate this mission to reach out to grassroots level. Finally, it is concluded that ICT is a significant contributor to the success of the ongoing initiative of Startup India.

In order to recognize and reward the extraordinary performance and achievements by ICT and allied sectors and promote universities, researchers and students through their research work adapting new scientific technologies and innovations. The two days conference had presentations from the researchers, scientists, aademia and students on the research work carried out by them in different sectors.

ICT4SD Summit is a flagship event of G R Foundation. This is the fourth edition. The Summit was addressed by eminent dignitaries including Shri Manguirsh Pai Raikar, Chairperson, ASSOCHAM MSME National Council; and Shri. Prajyot Mainkar, Chairman, IT Committee of Goa Chamber of Commerce and Industry; Mike Hinchey, President, IFIP and Chair IEEE, UK and Ireland; Milan Tuba, Vice Rector for International Relations, Singidunum University, Serbia; Prof Lance Fung, Australia; Prof. Jagdish Bansal, India; Mr. Aninda Bose, Springer; and Dr. Amit Joshi, Director, G R Foundation.

The overall conference had one Inaugural Session, one Keynote Session and 18 technical sessions during two days.

Zhuhai, Macao — Simon Fong
Kolkata, India — Nilanjan Dey
Ahmedabad, India — Amit Joshi

Contents

Editors and Contributors

About the Editors

Simon Fong graduated from La Trobe University, Australia, B.Eng. Computer Systems degree and a Ph.D. Computer Science degree in 1993 and 1998, respectively. Simon is now working as an Associate Professor at the University of Macau, China. Prior to his academic career, Simon held various managerial and technical positions, such as systems engineer, IT consultant, and e-commerce director in Australia and Asia. Dr. Fong has published over 300 international conference and peer-reviewed journal papers, mostly in the areas of data mining and optimization algorithms.

Nilanjan Dey is an Assistant Professor at the Department of Information Technology, Techno India College of Technology, India. He has authored/edited more than 75 books with Springer, Elsevier, Wiley, and CRC Press and published more than 300 peer-reviewed research papers. Dr. Dey is the Editor-in-Chief of the International Journal of Ambient Computing and Intelligence; a Series Co-Editor of Springer Tracts in Nature-Inspired Computing (STNIC); and a Series Co-Editor of Advances in Ubiquitous Sensing Applications for Healthcare, Elsevier.

Amit Joshi is currently the Director of Global Knowledge Research Foundation, also an Entrepreneur & Researcher who has completed his Masters and research in the areas of cloud computing and cryptography in medical imaging. Dr. Joshi has an experience of around 10 years in academic and industry in prestigious organizations. Dr. Joshi is an active member of ACM, IEEE, CSI, AMIE, IACSIT, Singapore, IDES, ACEEE, NPA, and many other professional societies. Currently, Dr. Joshi is the International Chair of InterYIT at International Federation of Information Processing (IFIP, Austria). He has presented and published more than 50 papers in national and international journals/conferences of IEEE and ACM. Dr. Joshi has also edited more than 40 books which are published by Springer, ACM, and other reputed publishers. Dr. Joshi has also organized more than

50 national and international conferences and programs in association with ACM, Springer, and IEEE to name a few across different countries including India, UK, Europe, USA, Canada, Thailand, Egypt, and many more.

Contributors

Almoataz Y. Abdelaziz Faculty of Engineering & Technology, Future University in Egypt, Cairo, Egypt

J. Abhiram KLE Technological University, Hubli, Karnataka, India

Adnan M. Abu-Mahfouz Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa

Vidhi Agarwal Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India

Istiaq Ahmed Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

A. Akshata School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

Redwanul Alam Alif Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Fahiem Altaf Department of Information Technology, Indian Institute of Information Technology Allahabad, Prayagraj, India

S. K. Amarnath KLE Technological University, Hubli, Karnataka, India

Aditi Ambadkar School of Computer Engineering and Technology, MIT World Peace University, Pune, Maharashtra, India

Ania-Nin-Ania Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

R. A. Archana Department of CSE SVIT, VTU, Bengaluru, Karnataka, India

R. Arokia Paul Rajan Department of Computer Science, CHRIST (Deemed to be University), Bengaluru, India

E. Arul Department of Information Technology, Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India

Meenakshi S. Arya Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India

Md. Asadut Zaman Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Tafesse Asrat School of Electrical and Computing, Adama University of Science & Technology, Adama, Ethiopia;
(Center of Excellence) Transportation and Vehicle Engineering (CTVE) ADAMA, Adama, Ethiopia

Saif Ali Athyaab CBIT, Hyderabad, India

Mahmoud A. Attia Department of Electrical Power and Machines, Faculty of Engineering, Ain Shams University, Cairo, Egypt

Vasudha V. Ayyannawar Department of Computer Science and Engineering, Basaveshwara Engineering College (Autonomous), Bagalkot, Karnataka, India

Asifa Baba Islamic University of Science and Technology, Awantipora, Kashmir, India

Vishwanath P. Baligar K.L.E. Technological University, Hubballi, India

Radha Balki 1MIT Academy of Engineering, Alandi, Pune, India

Parashuram Bannigidad Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India

Ian Benitez Camarines Sur Polytechnic Colleges, Nabua, Camarines Sur, Philippines

Madhuri Bhalekar School of Computer Engineering and Technology, MIT World Peace University, Pune, Maharashtra, India

Ajit Bijapur School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

Niloy Biswas Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Prasad Borannavar Electronics & Radar Development Establishment, DRDO, Bangalore, KA, India

Imen Bouazzi Department of Industrial Engineering, King Khalid University, Abha, Saudi Arabia;
Electronic and Micro-Electronic Laboratory, FSM, University of Monastir, Monastir, Tunisia

Sidra Azmat Butt Information Systems Group, Tallinn University of Technology, Tallinn, Estonia

Venkamaraju Chakravaram Jindal Global Business School, O. P. Jindal Global University, Sonipat, Haryana, India

R. Arokia Priya Charles Vishwakarma Institute of Information Technology, Pune, India

H. Chetan CMR Institute of Technology, Bangalore, Karnataka, India

Chetana Chaudhari School of Computer Engineering and Technology, MIT World Peace University, Pune, Maharashtra, India

Neera Chaudhary Raj Kumar Goel Institute of Technology Ghaziabad, Uttar Pradesh, India

Mukesh Chavada Department of Information Technology, Sarvajanik College of Engineering & Technology, Surat, Gujarat, India

Manpreet Chavan Sardar Patel Institute of Technology, Mumbai, India

Ashok K. Chikaraddi K L E Technological University, Hubli, Karnataka, India

Farzana Noshin Choudhury Ahsanullah University of Science & Technology, Dhaka, Bangladesh

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

Omdev Dahiya University Institute of Engineering and Technology, Maharshi Dayanand University Rohtak, Haryana, India

Vipul Dalal Vidyalankar Institute of Technology, Mumbai, India

Asmita Deshpande Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India

Yash Deshpande Sardar Patel Institute of Technology, Mumbai, India

Jenish Dhanani S.V. National Institute of Technology, Surat, Gujarat, India

Puja Dhar Lovely Professional University, Phagwara, Punjab, India

M. R. Dileep Department of Master of Computer Applications, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India

Rashes Dobariya Department of Information Technology, Sarvajanik College of Engineering & Technology, Surat, Gujarat, India

Andrew Doria College of Computer Studies, Silliman University, Dumaguete City, Philippines

Vijay Kr. Dwivedi United College of Engineering and Research, Prayagraj, India

Youssou Faye University Assane Seck of Ziguinchor, Ziguinchor, Senegal

Vijay Kumar Garg Lovely Professional University, Phagwara, Punjab, India

Timothy Gargantiel College of Computer Studies, Silliman University, Dumaguete City, Philippines

Kristabel Gavin College of Computer Studies, Silliman University, Dumaguete City, Philippines

B. Geethanjali Department of Computer Science and Applications, Bangalore University, Bangalore, India

Jossy P. George CHRIST (Deemed to be University), Bengaluru, India

Shine K. George Union Christian College, Aluva, Kerala, India

Sneha Rose George Union Christian College, Aluva, Kerala, India

Pradeepini Gera Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India

Megha Ghadage School of Computer Engineering and Technology, MIT World Peace University, Pune, Maharashtra, India

Shraddha Ghogare CSE Department, University of Texas at Arlington, Arlington, TX, USA

H. M. Gireesha School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

Ulhaskumar M. Gokhale G.H. Raisoni College of Engineering and Technology, Nagpur, MH, India

Aniket D. Gundecha 1MIT Academy of Engineering, Alandi, Pune, India

Avadhesh Kumar Gupta School of Computing Science and Engineering, Galgotias University, Greater Noida, India

Dheeraj Gupta MIS, Eller College of Management, University of Arizona, Tucson, AZ, USA

Prateek Gupta Adobe Systems Noida, 2 ASET, Amity University, Noida, Uttar Pradesh, India;
Adobe, Ghaziabad, India

Sachi Gupta Raj Kumar Goel Institute of Technology Ghaziabad, Uttar Pradesh, India

Mohamed Hamdy Department of Electrical Power and Machines, Faculty of Engineering, Ain Shams University, Cairo, Egypt

Md. Mahmudul Haque Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

J. Harsh Raj CBIT, Hyderabad, India

Sanjeevakumar M. Hatture Basaveshwar Engineering College (Autonomous), Bagalkot, Karnataka, India

Ravindra S. Hegadi Department of Computer Science, Central University of Karnataka, Kalaburagi, India

Rajshekar Hiremath Department of Information Technology, Pune Institute of Computer Technology, Pune, India

James Hodson AI for Good Foundation, El Cerrito, USA

Mohammad Rezwanul Huq East West University, Aftabnagar, Dhaka, Bangladesh

Md. Imtiyaz Anwar Babasaheb Bhimrao Ambedkar University, Bihar, India

Rakibul Islam Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Sajid-ul Islam Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Nalini C. Iyer School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

M. L. Jadhav Department of E&TC Engineering, MET's Institute of Engineering, Nasik, M.S, India

Md. Rafsan Jamil Ahsanullah University of Science & Technology, Dhaka, Bangladesh

Vivaksha Jariwala Department of Information Technology, Sarvajanik College of Engineering & Technology, Surat, Gujarat, India

Sakshi Jha GGSIPU, Greater Noida, India

Vidhey Joshipura S.V. National Institute of Technology, Surat, Gujarat, India

Nagaveni Kadakol Basaveshwar Engineering College (Autonomous), Bagalkot, Karnataka, India

Diwyani Kaikade 1MIT Academy of Engineering, Alandi, Pune, India

D. Kalyani VNR Vignana Jyothi Institute of Engineering & Technology, Hyderabad, India

Pariza Kamboj Sarvajanik College of Engineering and Technology, Surat, India

Suvarna G. Kanakaraddi K L E Technological University, Hubli, Karnataka, India

Suma Karanth R. V. College of Engineering, Bangalore, India

Asif K. Karim Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India;
SIERS Research Laboratory, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India

Arun Khosla NIT, Jalandhar, Punjab, India

Reeta Koshy Sardar Patel Institute of Technology, Mumbai, India

Mythili Krishnan Accenture, Bangalore, India

Akash Kulkarni School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

Radhika V. Kulkarni Department of Information Technology, Pune Institute of Computer Technology, Pune, India

Amit Kumar Department of Computer Science & Engineering, SRK University, Bhopal, India

Sachin Kumar Department of Electrical Engineering, IIT (BHU), Varanasi, Uttar Pradesh, India

Suresh Kumar Ambedkar Institute of Advanced Communication Technologies & Research Government of NCT of Delhi, New Delhi, India

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

Prashant Kushwaha Department of Information Technology, Indian Institute of Information Technology Allahabad, Prayagraj, India

Cat Kutay Charles Darwin University, Casuarina, NT, Australia

John Kyung College of Computer Studies, Silliman University, Dumaguete City, Philippines

Meghan Lendhe Sardar Patel Institute of Technology, Mumbai, India

Syeda Saiara Lubaba Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Manish Mahajan Department of Information Technology, Pune Institute of Computer Technology, Pune, India

Manu Mahajan SIRD, HIPA, Shimla, Himachal Pradesh, India

Khan Raqib Mahmud University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Shakil Mahmud East West University, Aftabnagar, Dhaka, Bangladesh

Soumyadev Maity Department of Information Technology, Indian Institute of Information Technology Allahabad, Prayagraj, India

Moname Majumder East West University, Aftabnagar, Dhaka, Bangladesh

Komal Malshikare Department of Information Technology, Pune Institute of Computer Technology, Pune, India

H. R. Mamatha PES University, Bangalore, India

Venkatesh Mane School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

T. N. Manjunath Department of ISE, BMS Institute of Technology and Management, VTU, Karnataka, Bengaluru, India

Nafees Mansoor University of Liberal Arts Bangladesh (ULAB), Dhaka, Bangladesh

Elisha Didam Markus Department of Electrical, Electronic and Computer Engineering, Central University of Technology, Bloemfontein, South Africa

Muthoni Masinde Department of Information Technology, Central University of Technology, Bloemfontein, South Africa

Siddhesh Masurkar Vidyalankar Institute of Technology, Mumbai, India

Chriss Acca Mathews Accenture, Bangalore, India

Fahad Mehfooz GGSIPU, Greater Noida, India

Ritika Mehra DIT University, Dehradun, Uttarakhand, India

Jinam Pankajbhai Mehta Department of Information Technology, Sarvajanik College of Engineering & Technology, Surat, Gujarat, India

Rupa Mehta S.V. National Institute of Technology, Surat, Gujarat, India

Marius Dasylva Mendy University Assane Seck of Ziguinchor, Ziguinchor, Senegal

Soumya R. Metri Department of Computer Science and Engineering, Basaveshwara Engineering College (Autonomous), Bagalkot, Karnataka, India

Chuchi Montenegro College of Computer Studies, Silliman University, Dumaguete City, Philippines

A. H. Moon Islamic University of Science and Technology, Awantipora, Kashmir, India

Jenifa Mousumi East West University, Aftabnagar, Dhaka, Bangladesh

B. L. Muralidhara Department of Computer Science and Applications, Bangalore University, Bangalore, India

B. Sridhara Murthy Kakatiya Institute of Technology and Science, Warangal, India

Rajashekara Murthy S. R. V. College of Engineering, Bangalore, India

Meghana Nagesha CMR Institute of Technology, Bangalore, KA, India

Sachin Naik Symbiosis Institute of Computer Studies and Research, Symbiosis International Deemed University, Pune, India

Aayushi Nainan Student at Dhirubhai Ambani International School, Mumbai, India

Varsha Namdeo Department of Computer Science & Engineering, SRK University, Bhopal, India

Roshani Narkhede 1MIT Academy of Engineering, Alandi, Pune, India

A. V. Navaneeth Department of Master of Computer Applications, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India

M. Neha PES University, Bangalore, India

P. Neha School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

Abhishek A. Nimje MIT Academy of Engineering, Alandi (D), Maharashtra, India

S. R. Nirmala Gauhati University, Guwahati, Assam, India

P. C. Nissimagoudar School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

Mohammed Ahsan Raza Noori DIT University, Dehradun, Uttarakhand, India

Abdulla AL Nuaimi SecureTech LLC, Abu Dhabi, United Arab Emirates

ChukwuNonso Nwokoye Nigerian Correctional Service, Awka, Nigeria

Md. Omar Sharif Rajme University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Karin Oolu Information Systems Group, Tallinn University of Technology, Tallinn, Estonia

Obiajulu Ositanwosu Nnamdi Azikiwe University, Awka, Nigeria

E. Padmalatha CBIT, Hyderabad, India

Ingrid Pappel Information Systems Group, Tallinn University of Technology, Tallinn, Estonia

Satyen M. Parikh Faculty of Computer Applications, Ganpat University, Gujarat, India

Thaksen J. Parvat D. Y. Patil College of Engineering and Innovation Varale, Pune, India

Sirajum Monir Parvez Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Ajay M. Patel Faculty of Computer Applications, Ganpat University, Gujarat, India

Hiral R. Patel Faculty of Computer Applications, Ganpat University, Gujarat, India

Sanket Patel Department of Information Technology, Sarvajanik College of Engineering & Technology, Surat, Gujarat, India

Shreya Patel Sardar Patel Institute of Technology, Mumbai, India

Anupama V. Patil Dr. D. Y. Patil Institute of Engineering Management & Research, Pune, India

Atharva P. Patil MIT Academy of Engineering, Alandi (D), Maharashtra, India

Tanuja R. Patil K.L.E. Technological University, Hubballi, India

Bijan Paul University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Ambika Pawar CSE Department, Symbiosis International University, Pune, India

Suresh K. Peddoju Florida International University, Miami, USA

Govind Pole Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India

B. L. Pooja K L E Technological University, Hubli, Karnataka, India

S. Praveen CMR Institute of Technology, Bengaluru, Karnataka, India

T. Preeti K L E Technological University, Hubli, Karnataka, India

A. Punidha Department of Computer Science and Engineering, Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India

C. N. Pushpa Department of CSE, UVCE, Bangalore University, Bengaluru, India

K. R. Radhika Department of CSE, UVCE, Bangalore University, Bengaluru, India

Amzad Hossain Rafi Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Fahrin Rahman Islamic University of Technology, Dhaka, Bangladesh

Rashedur M. Rahman Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

K. Rajeshwar NIRD&PR, MoRD, GoI, Hyderabad, India

Syed Shahir Ahmed Rakin Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Masud Rana East West University, Aftabnagar, Dhaka, Bangladesh

Mohammad Rifat Ahmmad Rashid University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Davinder Singh Rathee School of Electrical and Computing, Adama University of Science & Technology, Adama, Ethiopia;
(Center of Excellence) Transportation and Vehicle Engineering (CTVE) ADAMA, Adama, Ethiopia

Sunitha Ratnakaram Jindal Global Business School, O. P. Jindal Global University, Sonipat, Haryana, India

Niloy Saha Roy Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh

Shanta Saha University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Samar Jyoti Saikia Gauhati University, Guwahati, Assam, India

S. Sailekhya Prowess Enterprises, Vizag, India

Jatinderkumar R. Saini Symbiosis Institute of Computer Studies and Research, Symbiosis International Deemed University, Pune, India

Shreya Saini GGSIPU, Greater Noida, India

R. K. Saket Department of Electrical Engineering, IIT (BHU), Varanasi, Uttar Pradesh, India

Dipti Y. Sakhare MIT Academy of Engineering, Alandi (D), Maharashtra, India

V. M. Sardar Department of E&TC Engineering, J.S. College of Engineering, Hadapsar, Pune, M.S, India

Kumari Sarita Department of Electrical Engineering, IIT (BHU), Varanasi, Uttar Pradesh, India

Neerajkumar S. Sathawane G. H. Raisoni College of Engineering, Nagpur, MH, India

Philip Alger Serrano Camarines Sur Polytechnic Colleges, Nabua, Camarines Sur, Philippines

M. Z. Shaikh Shree Bhagubhai Mafatlal Polytechnic, Mumbai, M.S, India

G. A. Shanmugha Sundaram Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India;
SIERS Research Laboratory, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India

Dhananjay Kr. Sharma United Institute of Technology, Prayagraj, India

Garima Sharma Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

Nidhi Sharma GGSIPU, Greater Noida, India

Rishibha Sharma Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India

Sushma Sharma Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India

R. M. Shet School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India

Raghavendra M. Shet KLE Technological University, Hubli, Karnataka, India

Adapa Shivani PES University, Bangalore, India

Amala Siby CHRIST (Deemed to be University), Bengaluru, India

S. Shreyas CMR Institute of Technology, Bengaluru, Karnataka, India

M. R. Shrihari Department of ISE, BMS Institute of Technology and Management, VTU, Karnataka, Bengaluru, India

Dhirendra Kr. Shukla MNNIT Allahabad, Prayagraj, India

Archana Singh Adobe Systems Noida, 2 ASET, Amity University, Noida, Uttar Pradesh, India;
ASET, Amity University, Noida, UP, India

Sahil Singh GGSIPU, Greater Noida, India

Samridhi Singh CMR Institute of Technology, Bengaluru, Karnataka, India

M. SivakKumar Department of Computer Science & Engineering, SRK University, Bhopal, India

Kamna Solanki University Institute of Engineering and Technology, Maharshi Dayanand University Rohtak, Haryana, India

Harshita Sonkar Department of Information Technology, Indian Institute of Information Technology Allahabad, Prayagraj, India

Andy Spezzatti UC Berkeley, Berkeley, USA

Madhan Kumar Srinivasan Accenture, Bangalore, India

P. Sujith Kumar Union Christian College, Aluva, Kerala, India

Nancy Sukhadia Sarvajanik College of Engineering and Technology, Surat, India

Anika Tabassum East West University, Aftabnagar, Dhaka, Bangladesh

Neelakantam Tatikonda Wolkite University, Wolkite, Ethiopia

Manav A. Thakur D. Y. Patil College of Engineering and Innovation Varale, Pune, India

Naledi Portia Thothela Department of Information Technology, Central University of Technology, Bloemfontein, South Africa

J. Thriveni Department of CSE, UVCE, Bangalore University, Bengaluru, India

Vikas Tripathi Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

Anya Trivedi S.V. National Institute of Technology, Surat, Gujarat, India

Aashka Trivedi S.V. National Institute of Technology, Surat, Gujarat, India

Munesh C. Trivedi National Institute of Technology, Agartala, India

Ikechukwu Umeh Nnamdi Azikiwe University, Awka, Nigeria

R. Urvi CMR Institute of Technology, Bengaluru, Karnataka, India

Anagha Vaidya Symbiosis Institute of Computer Studies and Research, Symbiosis International Deemed University, Pune, India

Sourabh Varshney S.V. National Institute of Technology, Surat, Gujarat, India

Abhijith Venugopal PES University, Bangalore, India

K. R. Venugopal Department of CSE, UVCE, Bangalore University, Bengaluru, India

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

G. Vidyasagar Rao UCCBM, Osmania University, Hyderabad, India

P. Vijay Kumar VNR Vignana Jyothi Institute of Engineering & Technology, Hyderabad, India

Vishal S. Walunj CSMU, Panvel, India

Fahim Rubaiyat Zahir East West University, Aftabnagar, Dhaka, Bangladesh

Monji Zaidi Electronic and Micro-Electronic Laboratory, FSM, University of Monastir, Monastir, Tunisia;
Department of Electrical Engineering, King Khalid University, Abha, Saudi Arabia

Analysis and Development of Zero-Emission Individual Transportation Vehicles—E-bike

Davinder Singh Rathee and Tafesse Asrat

Abstract Urban development in Ethiopia has resulted in a rise in the number of cities, along with the indicated concentration of population in cities and existing urban agglomerations. So rise in demand for vehicles is also obvious for intercity or intracity mobility of goods and passengers. The present work deals with the activity carried out to develop transport for individuals and physically challenged people at a minimum price without the necessity of licence, which is an electric bicycle (E-bike). The integration of traditional bicycle with DC gear motor, controller, etc., to help in propelling known as E-bike. The main factors responsible for power consumption in propelling a bicycle are rider's weight, cargo weight if any, wind resistance, the frictional resistance of tire with surface and drag coefficient of paths as well as elevation angle of the path at normal speed. By considering all these parameters, total resistance faced by bike and torque required to pull the wheel has been calculated from the mathematical model specifically for local surrounding. The electric propulsion required to assist the rider to carry 92 kg of gross mass at the maximum speed of 25 km/h with 10° of the elevated path has been calculated approx 213 W and verified with simulation results carried out in MATLAB/Simulink platform. In addition to that basic model for the controller circuit been designed using Proteus and implemented using locally available components in the market and also has been tested on the bench test for 5 h with no load.

Keywords E-bike · Green revolution · Intercity transportation · RPM · Rolling resistance · Torque Proteus · Controller · Batteries

D. S. Rathee (✉) · T. Asrat
School of Electrical and Computing, Adama University of Science & Technology, Adama, Ethiopia
e-mail: jerryrathee@gmail.com

(Center of Excellence) Transportation and Vehicle Engineering (CTVE) ADAMA, Adama, Ethiopia

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_1

1 Introduction

Electric vehicles especially E-bike is the fastest growing sector in the automobile industry due to manifold benefits: firstly, zero-emission as well as health and secondly alternative individual transport mechanism for intercity mobility, without a licence, parking fees, insurance fees and registration fees [1, 2]. E-bikes look likes similar to traditional ancients bicycle just by adding a few electrical components like a controller, rechargeable battery and DC gear motor to enhance the human mobility for short distances.

The cyclists can choose support from motor optionally either for an elevated path to get a boost or to cruise over rough terrain at a specific speed maximum up to 25 km/h to travel the distance sooner. After that, motor stops assisting automatically and rests managed by cyclist pedalling. In Europe places like Birling and Amsterdam, E-bikes are assisting senior citizens to continue their ridings for coming many more years [3]. In addition to that users have reported miscellaneous benefits such as achieving high speed with minimum efforts [4], efficient time management, less challenging while elevated path, i.e. hills, etc. [5], as well as a lot of health benefits due to cycling. Moreover, most importantly, it does not cause noise and air pollution which supports the burning demand of today, i.e. green revolution.

All these advantages attract the concentration of the world, and it has been noticed that huge demand for E-bike is raised, i.e. China leads the world in sales, followed by the Netherlands and Germany. China manufacturing 36 million E-bike yearly [6], and exponential growth in sales of E-bike has been recorded in Europe especially 44% leading by Germany and 21%, respectively, followed by Netherlands [7]. Also, the number of E-bike sold in the UK has grown up by the number of 50,000 per year though it is not a major mode of transport [8] and seems to continue further. The survey study declares that 5% of UK adults are likely to buy an E-bike in the next 12 months [9]. There may the number of reasons that cyclist chooses to get assistance for pedalling such as carrying luggage, hilly routes, recovering from illness or injuries and the aged person who wishes to cycle to work [10].

2 Fundamentals of Electric Propulsion

The major component of E-bike is an electric motor, battery pack, charger and controllers as shown in Fig. 1. In this paper, two Li-ion 12 V/2A batteries are used as a power source and BLDC 250 W motor with a gear ratio of 8 and with a transmission efficiency of 0.9 is used to propel the bike and rider. The speed can be varied by mechanical force on the pedal by rider and throttle control. The throttle is variable resistance whose value decreases with the clockwise movement of the throttle. Electric propulsion replaces or supplements muscle power. This adds to the bicycle an electric motor, gear reducer, battery and power control. The speed of the motor will

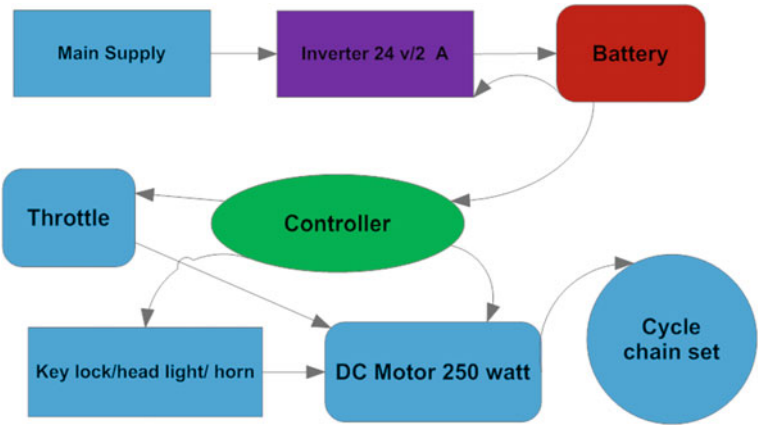


Fig. 1 Block diagram of main components of E-bike

be varied according to the throttle position. The output of throttle fed to the motor through the driver circuit.

To describe the behaviour of electric propulsion, some mechanical and mathematical knowledge is required. An electric bicycle has a conventional bicycle frame, pedals, cranks, chain and freewheel assembly. When a bike moves forward, following parameters $R =$ radius of the tires, $r_1 =$ radius of the sprocket, $r_2 =$ radius of the chainring and $\omega =$ angular velocity of the tires play a significant role. Also when a force $F =$ applied vertically to pedal arm, the reactive push backwards is generated given by F_1 and let F_2 is a forward force due to applied torque (Fig. 2).

$F = Ma$ (where $a = aR$ angular acceleration) and $F = F_2 - F_1 = MRa$

Torque (t) $= Ia$ (where I is the moment of inertia of the back tire)

$F_1R = Ia$ because the moment of inertia of both wheels is roughly the same

$Fp(r_1/r_2) = a(2I + MR2)$ by putting $a = Ra$ the linear acceleration can be given as

$$a = RFp(r_1/r_2)$$
$$(2I + MR2)$$

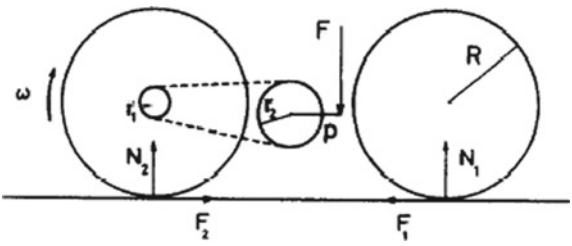


Fig. 2 Mathematical model for E-bike

Table 1 Lists the range of mass that is considered to be propelled

Component	Mass in kg
Bicycle frame	10–14
Motor and gear	8–10
Controller	1/2–1
Battery	4–7
Cyclist	60–90
Total weight	92.5–122

From the above equation, the smallest the gear ratio the largest the acceleration. Now let the angular velocity of the sprocket ($r_1 = w_1 = w$) and similarly angular velocity of the chainring ($r_2 = w_2$). The velocity due to cyclist's foot has given at pedal may be written by vapp = pw2.

Then velocity of bike can be written as vb = Rw = Rw1 = $w_2(r_2/r_1)R$

Substituting vapp/p for w_2 and rearranging, we obtained that vb/vapp = r_2R/r_1p and

$$v_b/v_{\text{app}} = (r_2/r_1)R/p$$

Also from $F = ma$ larger, the ratio produces more force. So it is also clear that distance travelled = circumference of the equivalent wheel, and it is 56 inch in this paper. Other related parameters which are considered for paper are given in Table 1.

The bicycle assembly is approx 14 kg with a 28 cm tire diameter available in the market. Also, the weight for 250 W motor with gear is the maximum up to 14 kg with a 7 kg battery. In addition to that other important parameters like rider's weight 75 kg, cargo weight 15 kg, rolling resistance 0.02, the drag coefficient of path 0.0032, the air resistance 4.16 N as well as elevation angle of the path 10° at speed of 25 km/h, it has been found that power required to propel is 253 W. Headwind speed = 25 km/h, so the power required to propel the bicycle and rider for specific environmental conditions is 253 W. By considering all the above parameters, result obtained from MATLAB platform has been presented in Fig. 3 and found in well agreement with mathematical model and calculations.

Linear distance travelled = 2 × 3.14 × 0.280 = 1.7584 m

Speed = 25,000/3600 = 6.94444 m/s

RPM = 25,000/1.7584 = 237

Total resistance = (rolling resistance + air resistance + track coefficient)

Power = [(M in kg) × (acceleration due to gravity) × (velocity m/s) × Rr] + [air density × Cd × area m^2 × v_3]

Total resistance is the summation of rolling resistance of vehicles, the air resistance of vehicles as well as gradient resistance for the elevated path. When E-bike goes up on the inclined road (0°–60° max), the total resistance faced also up as shown in Fig. 3a. Figure 3b also clears that torque required is directly portioned to the elevated path. Figure 3d has been observed that to propel the E-bike at the speed of 25 km/h at 10° of elevation, the required motor power is approx 250 W.

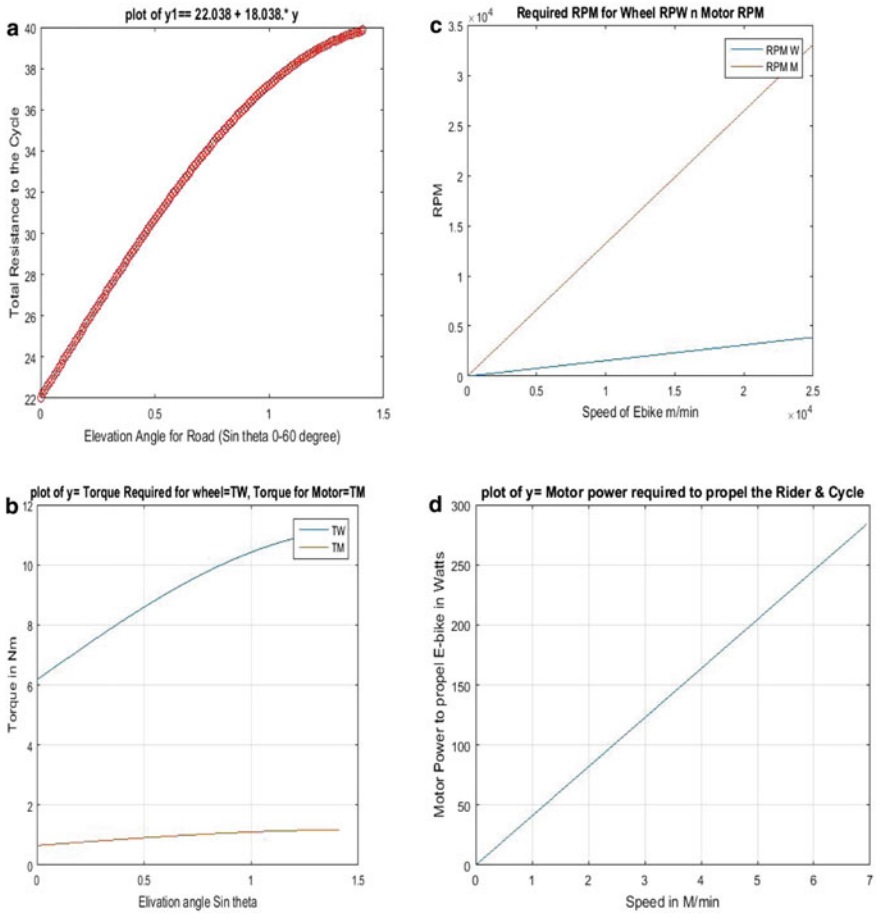


Fig. 3 **a** Total resistance faced. **b** The torque required to propel the E-bike. **c** Revolution per minute (RPM). **d** Motor power needed to propel the E-bike

In addition to that, we have developed the controller circuit using Proteus and implemented using locally available components. Further, this circuit has been tested for bench test successfully for 5 h under a no-load condition with a 250 W motor. The main function of the controller is sensing all the inputs from electric component, i.e. throttle, motor, brake, On-Off switch, sensors, etc., and then signalled the motor and light accordingly. The fundamental block for any controller is shown in Fig. 4. The other safety functional also possible, i.e. over and under current and voltage protection, temperature control protection as well as break protection, depends on make and design. For example, if the user applies brake and throttle at the same time, the brake function must stop the motor for safety.

The basic controller circuit is designed using Protus7 as shown in Fig. 4 and implemented using basic components. When key switch is on, it will connect the

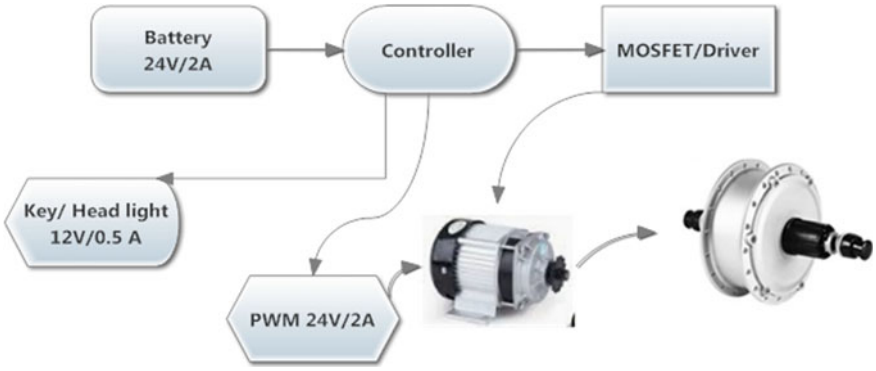


Fig. 4 Basic controller functions in E-bike

battery with the motor. The duty cycle of the signal will be assigned by IC 555 by comparing the threshold voltage from throttle input. The signal will be assigned to the MOSFET driver circuit which intensity will be similar to the position of the throttle. A controller is composed of resistors, capacitors, MOSFET, diodes, ICs, etc. The main circuit in a controller is the pulse width modulator (PWM) generator circuit to the MOSFET drive circuit, which controls the speed of a motor via changing the duty cycle of the pulse by sensing the movement of throttle (accelerator) as shown in Fig. 5. The controller supplies the required working voltages to all other subunits key lock, horn, brake light, sensors if any, motor, etc.

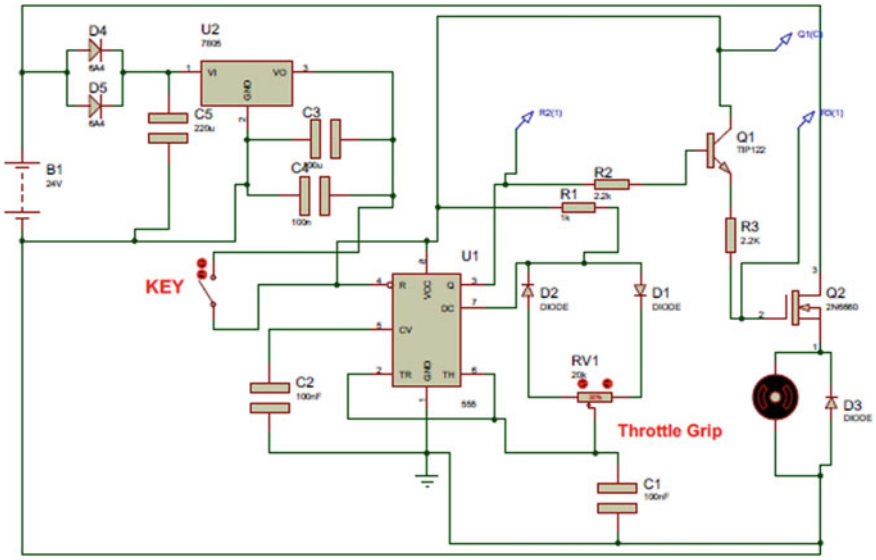


Fig. 5 Controller circuit diagram for E-bike

3 Conclusion

Nowadays, the ICE-3 wheeled vehicles are extensively used in almost all cities in Ethiopia. So considering the futuristic demand of community, i.e. environment friendly, affordable, independent, maintenance-free and does not require any licence as well, can be modified easily for specially challenged people.

This is an innovative start-up in the direction of developing the prototype for electric vehicles (EV). Especially in this work, our aim is to convert the traditional cycle into E-bike at Adama University for various benefits like supporting the green revolution, saving foreign currency, quieter in operation, cheaper for transportation, lowering service and maintenance cost when compared to motorized vehicles. The analysis presented in this paper can help the electric vehicle's designer to calculate the electric propulsion required for E-bike and then further extended for tricycle and electric three-wheelers and electric cars. Further, sensors can be included to provide circuit protection. The system performance can be improved by applying solar power system.

References

1. Cong, R., Martinez, R., Casilang, M., & Vong, P. (2010). *Electric bicycle system.* Senior Project California Polytechnic State University, San Luis Obispo.
2. http://www.Electric-bicycle-guide.com/electric bicycle history.html.
3. Andrew, J. (2019). Hawkins@andyjayhawk, Ditch your car, delete your Uber app, skip the subway—now is the time to get an electric bike.
4. Cycling Weekly. (2018). *Seven of the best electric bikes for 2018: All you need to know about e-bikes.* https://www.cyclingweekly.com/group-tests/best-electric-bikes-need-know-e-bikes-322613. Date Accessed August 10, 2018.
5. Transport for London. (undated). *What is an e-bike?* https://tfl.gov.uk/modes/cycling/start-cycling/electric-bikes/what-is-an-e-bike. Date Accessed September 7, 2018.
6. Davies, A. (2017). *Essential evidence on a page: The rise and rise of e-bikes.* https://travelwest.info/project/ee-164-rise-rise-e-bikes. Date Accessed September 5, 2018.
7. Fishman, E., & Cherry, C. (2016). E-bikes in the mainstream: Reviewing a decade of research. *Transport Reviews, 36*(1), 72–91.
8. BBC. (2018). *Could e-bikes be the future?* https://www.bbc.co.uk/news/uk-england-london-43836937. Date Accessed August 10, 2018.
9. Mintel (2018) cited in iNews. (2018). *Why electric bikes have become so popular*. https://inews.co.uk/news/health/why-electric-bikes-have-become-so-popular/. Date Accessed August 31, 2018.
10. Halfords cited in iNews. (2018). *Why electric bikes have become so popular*. https://inews.co.uk/news/health/why-electric-bikes-have-become-so-popular/. Date Accessed August 31, 2018.
11. www.rospa.com.
12. www.slideshare.net.
13. www.ti.com.
14. www.tandfonline.com.
15. www.sewanee.edu.

Development of a Flood Hazard Prediction Model Using Artificial Neural Network

Chuchi Montenegro, Andrew Doria, Timothy Gargantiel, Kristabel Gavin, and John Kyung

Abstract Flooding has become a major problem in most cities, and the need of a system that maps flood hazard is extremely useful to assess potential consequences, mitigate impact of large flood events, disaster preparedness, post-flooding relief operations, and emergency response team planning. This work describes the development of a machine learning prediction model that can be used to predict the severity of flood within Dumaguete City. The model was created using historical atmospheric, meteorological, flood, ground elevation, topographic, and geological data. Two preliminary prediction models were created using feedforward neural network and self-organizing map to map areas vulnerable to flooding within Dumaguete City. Fifty (50) different scenarios of selected dates from 2000 to 2014, the years where extreme flooding was experienced in the city, were tested to the two predictive models to reproduce flooded areas for verification. Observation on the tests performed on both models results to 80% and 86% prediction rate, respectively.

Keywords Flood hazard mapping · Neural network · Prediction model

1 Introduction

Flooding is a major problem worldwide and is one of the most frequent, disastrous, and widespread natural hazards [1]. Not only it a serious threat to life but also to property and livelihood. Floods have been a part of Filipino history as the country is located in the path where mostly torrential monsoon rains and tropical cyclones are moving. According to the record of International Emergency Disaster Database (EM-DAT), there are 37 occurrences of flood in the Philippines from 2010 to 2013. The floods caused by tropical cyclone Washi (local name Sendong) killed 1268 Filipinos, mostly in Cagayan de Oro and Iligan City when it hit the Philippines in December 2011. In 2012 alone, there had been 143 reported incidents of floods which affected 7,287,951 people [2], and in January 2014, almost 900,000 have been affected by

C. Montenegro (✉) · A. Doria · T. Gargantiel · K. Gavin · J. Kyung
College of Computer Studies, Silliman University, Dumaguete City, Philippines
e-mail: chuchi.montenegro@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_3

tropical depression Agaton. According to authorities, 45 were dead while 260,000 were displaced in the southern Philippines mostly in Caraga region and Davao [3].

Dumaguete City is a small coastal city in Negros Oriental, a province located in the region of Central Visayas, Philippines. It has a land area of 8310 acres and a population of over 131 thousand. The city has at most six kilometers of flat terrain from the coastline that gently slope upwards toward Mt. Talinis. It has two major rivers, the Banica River which runs through the city towards Mindanao Sea and Okoy River that partly passes on the northwest. When the city experienced nonstop rain for almost 24 hours in 2009, the two major rivers in the city overflowed causing floods, killed many lives, and damaged properties and crops. It was the first time the city experienced such catastrophe and was never prepared for it to happen. The incident was again followed by another flooding when typhoon Sendong which brought heavy rains in 2011 caused floods that swept throughout the city and neighboring towns killing 22 lives.

The damaged cost in the series of flooding in the city has caused thousands or not millions of money. Aside from the water damage to properties, there is a significant humanitarian cost in terms of both physical and mental health issues as a result of forced relocations and property damage. Given these facts, it is clear that there has been a need to develop improved methods for identifying the most cost-effective intervention strategies that best mitigate damage from flood events.

Advancements in computing technology have provided flood hazard mapping with powerful tools in modeling flood events, among them is artificial neural network (ANN), a mathematical model that allows computers to evolve behaviors based on empirical data. ANN has been applied in many case studies with different level of success, and despite the large amount of work published in this field, it is still a challenge to use ANN models reliably in different situations. Self-organizing map (SOM) is a type of ANN that is trained using unsupervised learning to produce a low-dimensional, discretized representation of the input space of the training samples, called a map. Feedforward neural network is the simplest type of ANN wherein connections between the nodes do not form a cycle. The objective of this thesis is to use SOM and FNN as prediction model on Geographic Information System (GIS) map to simulate severity of areas prone to flooding in Dumaguete City.

2 Methodology

This section describes the methodology of the work being performed. This is reflected in Fig. 1. In this study, three maps (geological data) from Dumaguete City Planning and Development (DCPD) office were utilized. These include road maps with built-up areas, a contour map, and map of rivers and waterways. The road map (Fig. 2) shows the major roads that include national, major, new, and diversion roads. The map also includes a convenient delineation of borders of all barangays in the city. The contour map (Fig. 3) shows contour lines. These contour lines are lines drawn

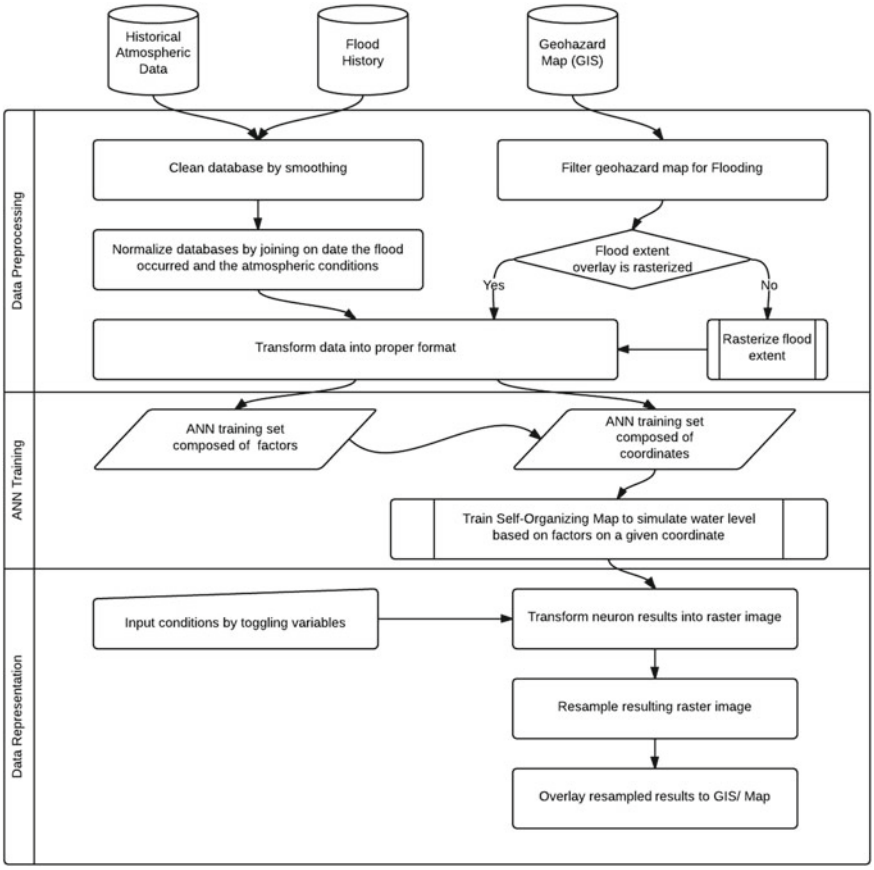


Fig. 1 Framework of the study

on a map connecting two equal elevations. It shows the elevation and the topography of Dumaguete and the shape of the terrain. It also includes data on buildup areas, represented in yellow color, which shows the areas where there are man-made infrastructures such as residential and commercial areas and which also shows where human population is concentrated. This data is important since these infrastructures slow the rate of movement and accumulation of flood but also prolong the duration of the flood to recede. The figure also shows the different barangays in the city and its respective border and is primarily used for referencing the coordinates of barangays and its arbitrary borders in accordance with the road map on Fig. 2. From the contour map, two separate maps were also generated. One is the map of the contours, and the other map is the built-up area map (Fig. 4). The rivers and waterways map show bodies of water in Dumaguete. As what is shown on Fig. 5, Banica River is a major natural spillway located in Negros Oriental. It starts from the Municipality of Valencia running down to the Negros Oriental's capital (Dumaguete

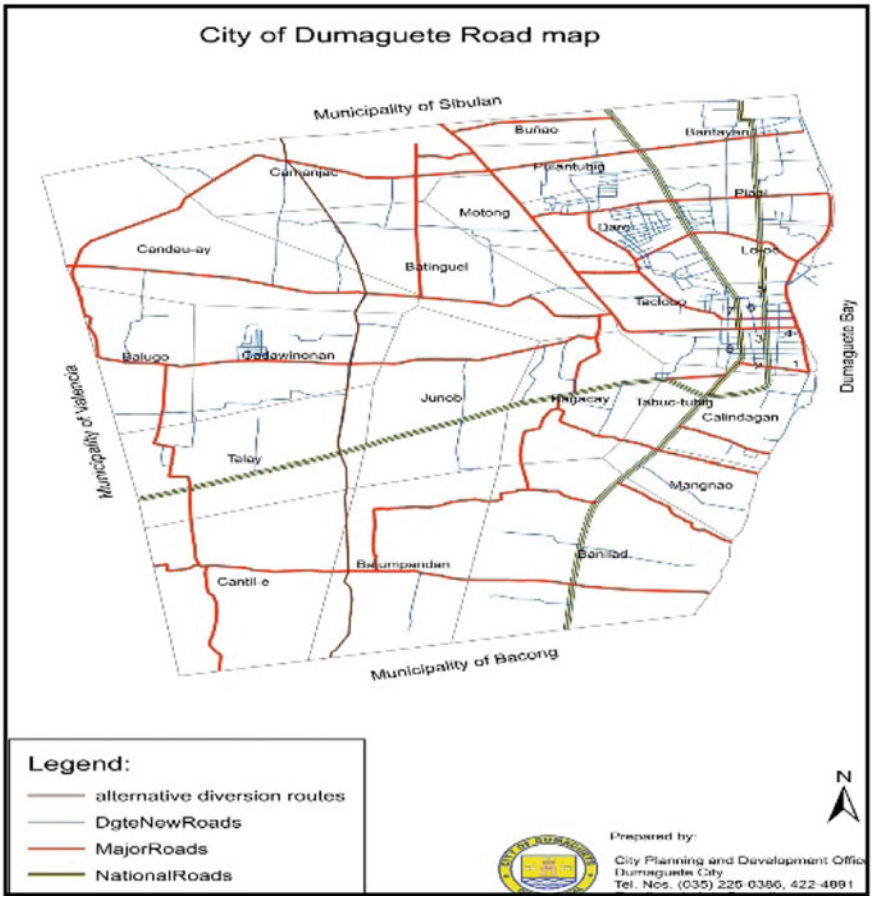


Fig. 2 Roadmap of Dumaguete City

City) through the 10 barangays. Okoy River is a major natural spillway located in Negros Oriental that starts from Municipality of Valencia running down to San Jose to barangay Candua-ay and ends in Sibulan.

Historical meteorological data of the city from the year 1975 to 2014 was gathered from Philippine Atmospheric Geophysical and Atmospherical Services Administration (PAG-ASA) while flood historical data from 1999 to 2013 was gathered from Disaster Risk Reduction and Management Office (DDRMO) of Dumaguete. This study also relied on Geographic Information System (GIS) in mapping and identifying topographic features and ground elevation. Point sampling algorithm, a plug-in tool of QGIS, was used in getting the elevation data which extracts pixel values at a certain location or aggregate them over the area.

Training of the ANN is performed by applying an optimizing algorithm, which attempts to reduce the error in the network output by adjusting the matrix of network

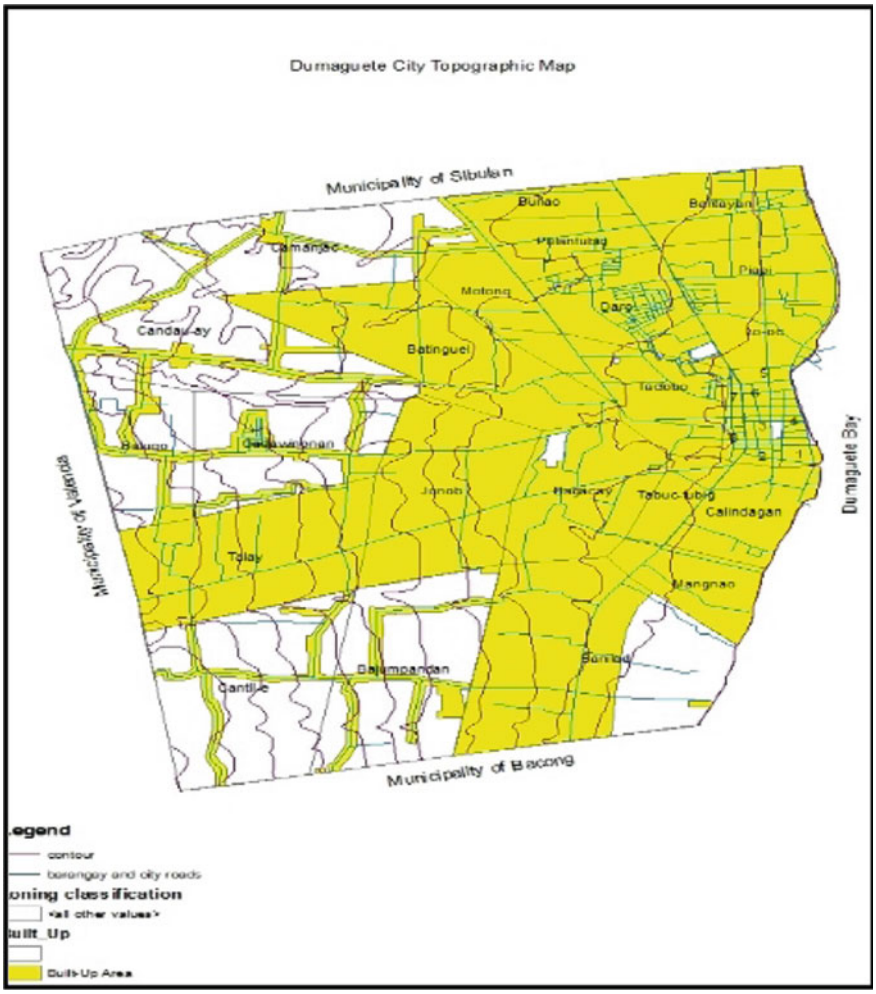


Fig. 3 Contour map of Dumaguete City

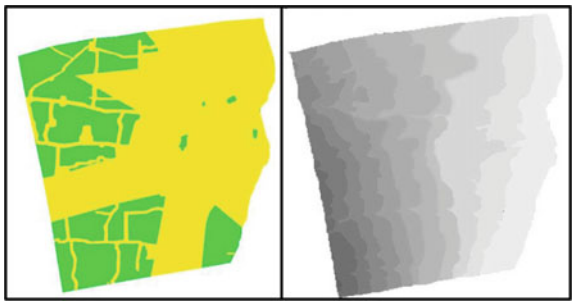


Fig. 4 Contour and buildup area map

Fig. 5 River and waterways map

weights and neuron biases. First, the gathered data is feed to the system consisting of different factors needed to generate GIS, soil data, weather data, topographic map, and drainage map. The data will then pass through a preprocessing tool to clusters the data and fed into the ANN which organizes the data further. Every data is fed into different units to perform a specific task. If the neural unit is not yet trained enough, it will go back to the previous step performing with another data of the same type. Since the unit receives specific task that they can do well enough, they become even more competent in that task. Then with the trained data units, the system can now generate a flood hazard map based on what the system has learned (Fig. 6).

There are two different outputs of the transformed data, a table representing the factors for flooding, and the coordinate locations and vulnerability. The first table represents the first layer while the second set represents the hidden layer. The final layer will be the competition layer. For easier understanding, training would go like this. The first layer would receive the conditions within the past two days. The conditions would be assessed whether the coordinates would experience flooding or not based on the geohazard map and to what depth in the second layer based on historical flood data. Finally, neurons would compete in the third layer to flood-prone areas in the resulting map.

Due to the large set of atmospheric and geographic data used as training sets and the limited hardware capacity and specifications used during training, only two networks were successfully modeled as epochs are adjusted to 25,000 level. FNN

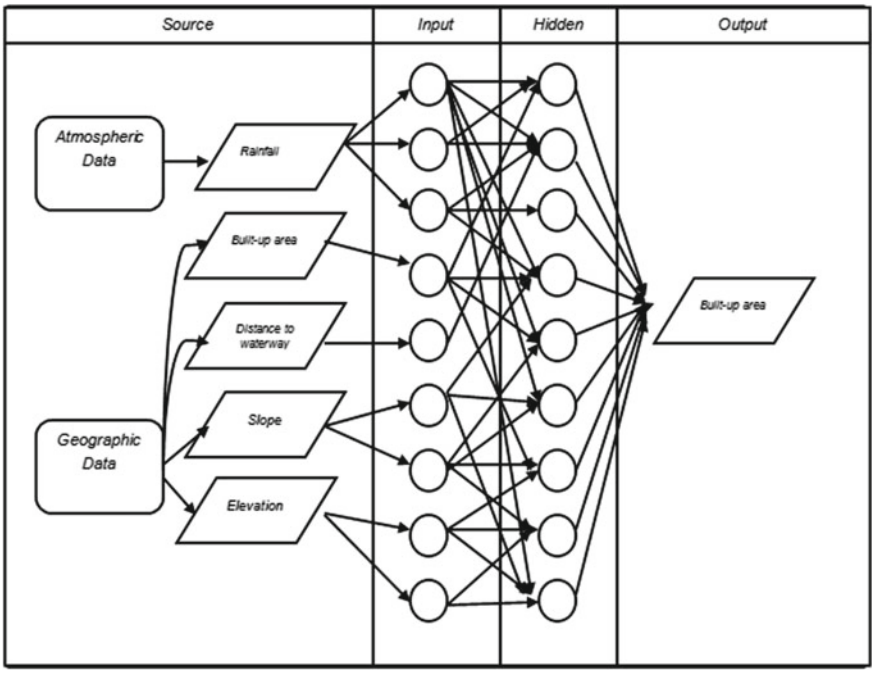


Fig. 6 Neural network model

or network 1 and SOM or network 2. As reflected in Figs. 7 and 8, error rate at the beginning was very high but acquires a steadier decline in error as the number of epochs increased. In this study, testing was performed on the two networks trained at 12,000 epochs because experiments of the epoch values show that at level 12,000, the error rate becomes steady.

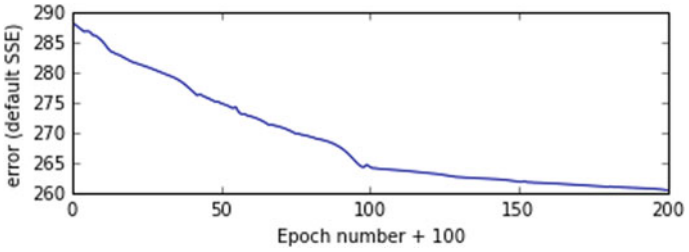


Fig. 7 Training of model 1

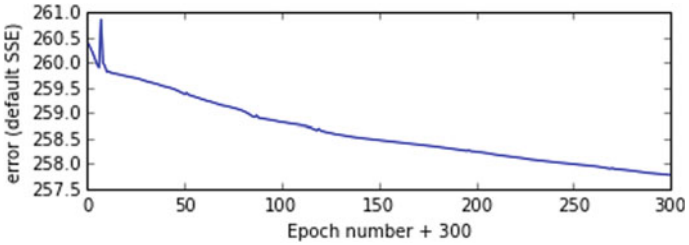


Fig. 8 Training of model 2

3 Results

To visualize flood severities in the map, the network model predicted value in the range of 0 to 1 is converted into color labels. The color labels are shown in Fig. 9 indicating red color as severe and green as normal.

The two networks generated after training were both tested on 50 different scenarios of selected dates from 2000 to 2014 (sample plotting as shown in Figs. 10 and 11) to test which between the two networks can predict flooding better. Although the results of the feedforward neural network and self-organizing map are similar, the results generated by network 1 when tested on the days when actual flooding occurred show less consistent results while network 2 shows flooded areas which are more or less equivalent to actual flooded areas especially on December 2011 and December 2014. Based on the data the researchers gathered from DRMMO and from records of barangays, flooded areas are usually limited near river banks which are better represented in the prediction of the second network. Observation on the tests

Fig. 9 Color label assignments for plotting

R	G	B	Weight
255	0	0	1 and above
255	51	0	.90 to .99
255	102	0	.80 to .89
255	153	0	.70 to .79
255	204	0	.60 to .69
255	255	0	.50 to .59
204	255	0	.40 to .49
153	255	0	.30 to .39
102	255	0	.20 to .29
51	255	0	.10 to .19
0	255	0	0 to .00
0	255	51	-.10 to -.01
0	255	102	-.20 to -.11
0	255	153	-.30 to -.21
0	255	204	-.40 to -.31
0	255	255	Below -.41

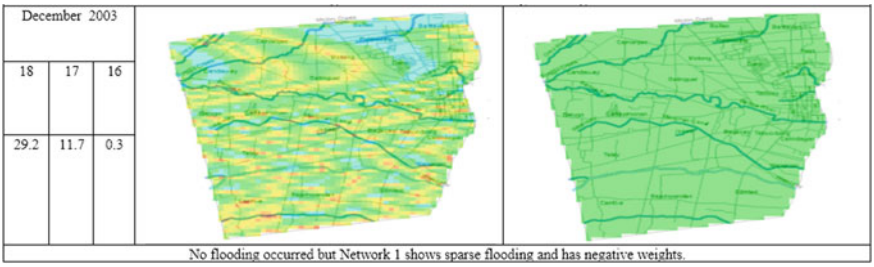


Fig. 10 December 2003 model test result of model 1 (left) and model 2 (right)

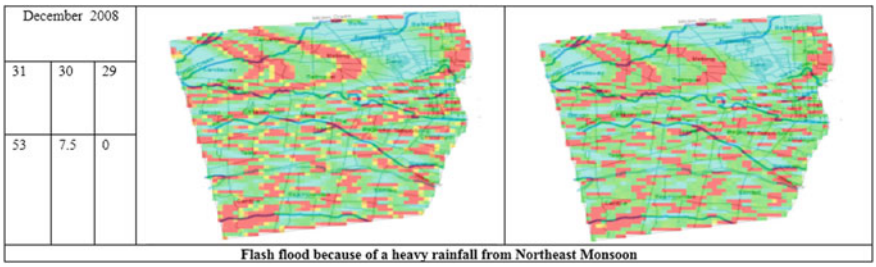


Fig. 11 December 2008 model test result of model 1 (left) and model 2 (right)

performed on both networks shows that the second network performed better at 86% prediction rate compared to 80% prediction rate of the first network. In this study, the accuracy of the result is computed based on the percentage of incorrect nodes from the total nodes of 2861.

4 Observation and Recommendation

Most of the area around Banica River was supposed to be exposed to flooding, but it was observed that there were gaps to the flood extend generated by the model but still congregate near the rivers. The area in barangay Cantil-e and Talay shows vulnerability to flood despite not being flood-prone albeit sparsely and severe. Since in this study there was manual plotting of distance between nodes, it is recommended to assign geographic points closer to each other to avoid gaps in the mapping. Also, due to the limitation of the hardware used in this research, the creation of the model was constraint to a limited epoch to a safe limit without resulting to memory failure; thus, it is suggested to use better hardware suited for handling training of large amounts of data.

References

1. National Institute of Geological Sciences, UP Diliman. (2012). *Philippine Flood Hazard Maps*. http://www.nababaha.como.
2. CDRC. (2012). *CDRC Database*. http://www.cdrc-phil.com/resources.
3. Chen, S., Jakeman, A., & Norton, S. (2008). Artificial intelligence techniques: An introduction to their use for modelling environmental systems. *Mathematics and Computers in Simulation, 78,* 379–400.
4. Department of Science and Technology. (2012). *Project NOAH*. http://noah.dost.gov.ph.
5. Kia, M. S., Pirastesh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. *An artificial network model for flood simulation using GIS: Johor River Basin, Malaysia.* http://www.environment.uwaterloo.ca/research/rsgtl/publications/Saied_Publications/6.pdf.
6. Kussul, N., Shelestov, A., & Skakun, S. *Intelligent Computations for Flood Monitoring*. http://sci-gems.math.bas.bg/jspui/bitstream/10525/1029/1/IBS-02-p06.pdf.
7. Macgillivray, K., Wilson, J. B., & Holt, A. (1999). *The Use of Artificial Intelligence Technique and GIS for Predictive Vegetation Modelling*.
8. Pan, T. Y., Lai, J. S., Chang, T. J., Chang, H. K., Chang, K. C., & Tan, Y. C. (2011). *Hybrid Neural Networks in Rainfall-Inundation Forecasting Based on a Synthetic Potential Inundation Database*. http://www.nat-hazards-earth-syst-sci.net/11/771/2011/nhess-11-771-2011.pdf.
9. Subramanian, R. (2012). Implementation of neural networks in flood forecasting. *International Journal of Scientific and Research Publications, 2*(10). http://www.ijsrp.org/research-paper-1012/ijsrp-p1096.pdf.
10. Zhou, C., Lou, J., Yang, C., Li, B., & Wang, S. *Flood Monitoring Using Multi-Temporal AVHRR and RADARSAT*. http://asprs.org/a/publications/pers/2000journal/may/2000_may_633-638.pdf.

Model-Based Design of Motion Control System

Roshani Narkhede, Radha Balki, Diwyani Kaikade, and Aniket D. Gundecha

Abstract Developers of complex control system aim to provide superior performance and compact designed equipment while considering the user requirements and keeping the costs down. One way to achieve higher performance and lower costs is by improving the approach toward design methodology of the complex systems. However, in today's world, the demand for systems having accurate and better performance is growing rapidly and reaching the point where traditional design and verification methodologies are falling apart. Thus, by using model-based design approach, engineers are able to test and find errors of an application at each stage and create a higher performance system. Model-based design optimizes the system behavior and quickly evaluates a variety of control strategies. It uses system simulation where engineers integrate software algorithms with the application. The model can be reused for real-time testing. This approach enables the engineers to simulate and test the hardware and software at the early stages of development cycle, and the errors that are found can be easily fixed and would be less expensive.

Keywords Control systems · Model-based design · Real-time testing · Development cycle · Verification · PID

R. Narkhede (✉) · R. Balki · D. Kaikade (✉) · A. D. Gundecha
1MIT Academy of Engineering, Alandi, Pune, India
e-mail: rnnarkhede@mitaoe.ac.in

D. Kaikade
e-mail: ddkaikade@mitaoe.ac.in

R. Balki
e-mail: rsbalki@mitaoe.ac.in

A. D. Gundecha
e-mail: adgundecha@entc.maepune.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_4

1 Introduction

In today's world, automation plays a very crucial role in each and every field. Modeling, simulation, control, verification and validation are the factors to be driven by a specified algorithm or model. Considering traditional workflow, it limits the engineers to test and validate a system at an earlier stage of development cycle. Thus, the errors/defects arose could often be resolved at the later stage during final system integration. Thus, this approach was suitable when behavior of the system was predictable. Model-based design alleviates these drawbacks by addressing it through simulation and up-front verification.

Model-based design has become a key workflow model for a wide range of motion control applications from industrial automation to consumer goods [1]. From simulation and early stage real-time testing during development cycle, model-based design provides flexibility to designers to create more robust, cost effective, and high-performance control systems for motion control applications [1]. Some or the other times control systems become progressively complex, thus, up-front verification before executing the hardware will not only be a good practice, it will be imperative.

2 Model-Based Design

See Fig. 1.

Executable Specifications from Models:

- Specifications and requirements are validated against user acceptance tests (UAT).

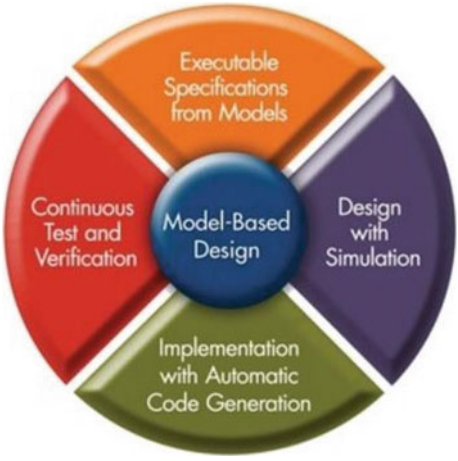


Fig. 1 Model-based design [2]

- Maximum design issues are identified through software and hardware analysis at this stage.

Design with Simulation:

- Design allows to check the functionalities, communication within the sub-models and outside world can be easily understood.
- Simulation helps designers and engineers check if the software and hardware are able to perform the assigned tasks. Code is integrated with the application.

Implement with Automatic Code Generation:

- The model is being simulated without any physical hardware components, i.e., model-in-loop testing is done before auto-code generation.
- Code is generated from the models using tools like Simulink coder, Ecoder, AUTOSAR, ERTL, etc.

Continuous Test and Verification:

- It involves unit testing, system testing, integration testing, and user acceptance testing, wherein the models bugs and errors are fixed after the code is integrated with the application.
- Test the complete application with its functionality, inter-dependency, and communication so that it meets the user's requirements.

3 Workflow

Workflow of Simulink based on model-based design (Fig. 2):

System definition and Component Analysis:

Identifying modeling objectives, components requirements, and layout of model [3].

Model analysis, simulation, and system verification:

Observing and evaluating the objectives of each development phase to check whether they meet determined requirements.

System Validation:

To test components, integrate algorithms with the model, test system [3]. To determine whether the model meets the end-user requirements.

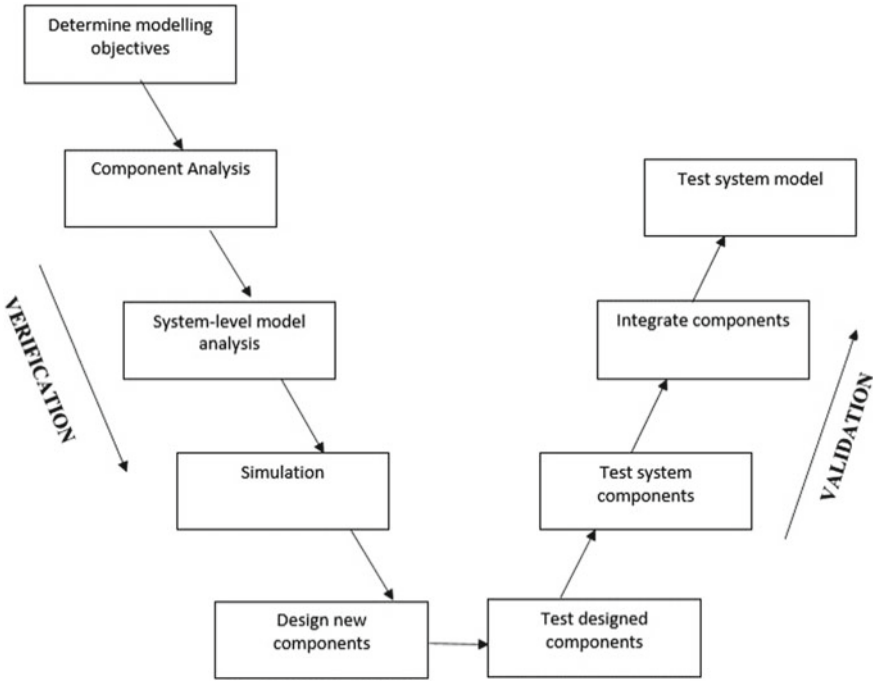


Fig. 2 Model-based design workflow in Simulink

4 System Identification and Modeling

4.1 System Identification

System identification is a process that utilizes systems input and output signals for the mathematical analysis of the system dynamics. System identification of DC motor by experimental analysis [4] (Table 1).

4.2 System Modeling

System modeling is a methodology where each model represents a different perspective of the system by creating conceptual models of the control systems.

A DC motor can be characterized as;

$$\omega(s)/E(s) = k/[(Js + b)(sL + R) + k2] \tag{1}$$

$$\theta(s)/E(s) = k/(s[(Js + b)(sL + R) + k2]) \tag{2}$$

Table 1 System identification of DC motor

Parameter	Symbol	Value
Motor armature resistance	R	4.86Ω
Motor armature inductance	L	22.248 mH
Motor voltage constant	K_e	0.01257 V/rpm
Motor torque constant	K_t	0.12003 N m/A
Motor inertia	J	4.7224×10^{-4} Kg m^2
Viscous friction coefficient	B	0.0828×10^{-3}N m s/rad
Coulomb friction coefficient	C	244.33×10^{-3} N m

where ω: angular speed;
θ: angular position;
E: applied voltage [4] (Fig. 3).

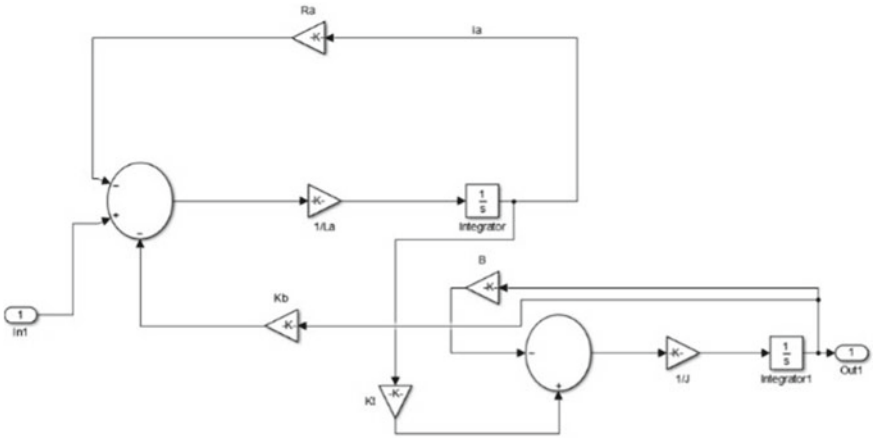


Fig. 3 Simulink model of motor

5 Input Test Signals to the System

5.1 Step Signal

A step signal is fed to observe the response of the system. Step signal though being a sudden input was preferred at the initial stage as the magnitude of the step signal does not changes with time.

5.2 Sinusoidal Signal

A sine wave when fed to the system leads to the verification, if the system follows the path of the fed signal or not.

5.3 Trajectory Generation

A trajectory is generated and fed to the system to ensure if the system executes smoothly along the sequence of points, i.e., the path.

5.3.1 The desired angular velocity trajectory ω^* is proposed as follows: The customized signal defines t is the time [5] (Fig. 4).

$$\omega^* = 2 + 1.75 * pi[(1 - e^{-2t^3})(1 - \sin(2.5t))] \tag{3}$$

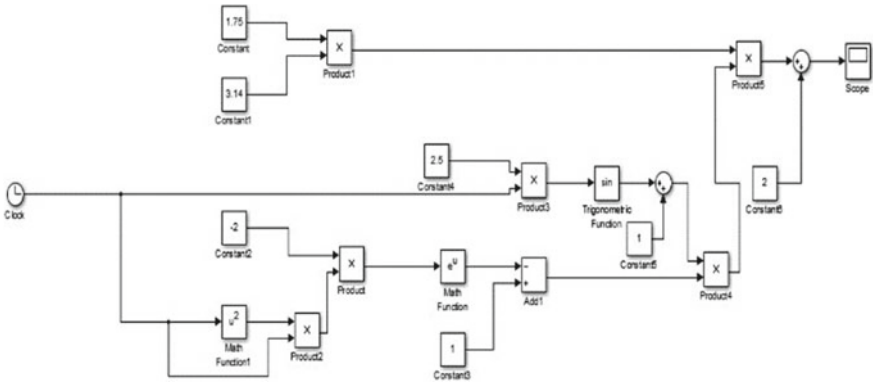


Fig. 4 Trajectory 1 → trajectory generation by Eq. 3

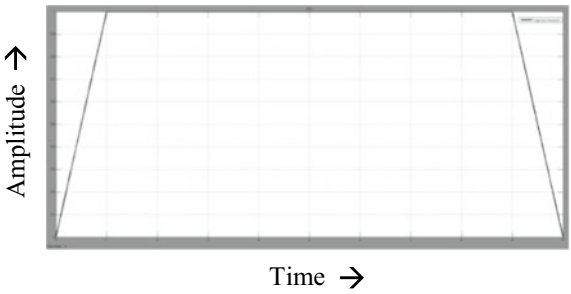


Fig. 5 Trajectory 2 → trajectory generation by Eq. 4

5.3.2 This trajectory is achieved so as to reduce errors and for an increased stable performance of the system. To reduce errors due to the jerks, the speed of the system is increased as well as decreased gradually once it achieves the maximum constant speed. Generation of a velocity profile is achieved by one of the equations of motion (Fig. 5),

$$V = U + AT \tag{4}$$

6 Proportional-Integral-Derivative (PID) Controller

PID controller can be termed as a widely utilized, understandable, moreover a simple, though a versatile feedback compensator structure [6]. Decoding about the controller, the integrator figures out the history of the system, conversely the differentiator predicts the future behavior of the system. We may term PID as the best-known controller used for wide industrial as well as academic purposes due to its simple as well as efficient structure. Setting of the PID parameters is termed as PID tuning which can be realized using various algorithms of which, Zeigler–Nichols tuning formula [7, 8] is a famous classical approach. PID controllers are designed for synchronization of time-domain behaviors of the different plants [9].

Features of P, I and D Controller

With increasing *KP*, proportionally increases control signal for the same level of error. The closed loop system tends to react more quickly, when applied to a given level of error; however, it also tends to overshoot more. An increasing *KP* reduces the steady-state error and the rise time with increasing overshoot.

The introduction of derivative term to the controller (*Kd*) enhances the potentiality of the controller to compute error. If the proportional control (*Kp*) is set, the only way to enhance the control is by increasing the error [6]. Keeping the magnitude of the error signal relatively small, if the error initiates sloping upwards, the control signal

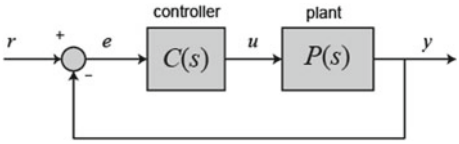


Fig. 6 Closed loop system with PID controller [6]

can become large when accompanied by the derivative control. This expectancy adds damping to the system, thereby decreasing the overshoot. Inclusion of *Kd* does not results in any changes in steady-state error.

The introduction of an integral term to the controller (*Ki*) assists the lessening effect of steady-state error. If there is a continuous steady-state error, the integrator assembles, thereby resulting in increased control signal and lowered error signal. A drawback of *Ki* is, it may result as a slow response system as well make it oscillatory once the error signal changes its sign.

As of an end result, one aims of a controller with fast rise time (*tr*), minimal overshoot along with zero steady-state error (*e*ss) [6] once we unite the proportional, derivative in addition to integral controller.

When implemented on the system, the P controller minimizes the steady-state error along with the rise time; shoots up the overshoot and a little bit decreases the settling time [9]. The PD controller reduces the *ts* and the overshoot and has insignificant change in *tr* and *e*ss. Bringing PI control along with helps eliminating the steady-state error. Putting together PID control, we observe no overshoot, eliminated *e*ss with a fast rise time. In a closed loop system, output of a PID controller = control input to the plant, i.e., calculated in the time-domain from the feedback error as [6]:

$$u(t) = K_p e(t) + K_i \int e(t)\mathrm{d}t + K_d \frac{\mathrm{d}e}{\mathrm{d}t} \tag{5}$$

where *r*—desired output, *y*—actual output, *e*—error (difference between desired output and actual output).

Figure 6 depicts the internal working of closed loop system which results in a transfer function. Taking Laplace transform of Eq. (5), the transfer function of the PID controller is obtained:

$$K_p + K_d s + \frac{K_i}{s} = \frac{K_p s + K_d s^2 + K_i}{s} \tag{6}$$

where K_p = proportional gain, K_i = integral gain, and K_d = derivative gain.

7 Results and Analysis

→ Output Waveform
→ Input Waveform

Closed Loop

In a closed loop control system, the output has an effect on the control action of the input, i.e., in the closed loop control systems, the output is feedback to the input. Therefore, they are termed as feedback control systems. In closed loop control systems, to get the output of the system at desired value the controller minimizes the actuating signal error produced due to disturbance (external or internal environment), i.e., the difference between the input and feedback signal of the system.

Closed loop without PID Controller

Closed loop system, i.e., the feedback system gives uncontrolled output response along with the path of the test input response (Figs. 7, 8, 9, and 10).

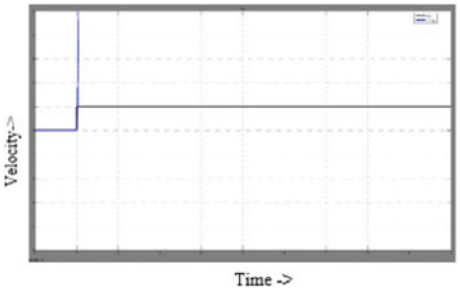


Fig. 7 Response of step input to the plant

Fig. 8 Response of sinusoidal input to the plant

Fig. 9 Response of trajectory 1 input to the plant

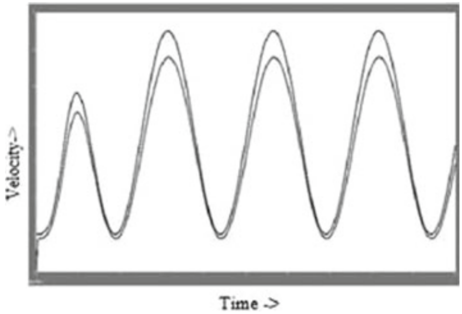
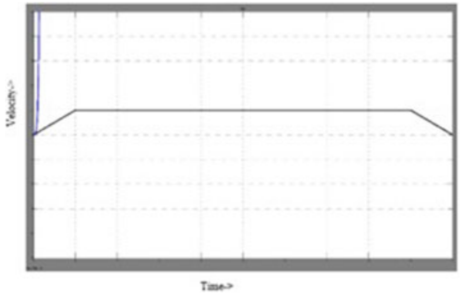


Fig. 10 Response of trajectory 2 input to the plant



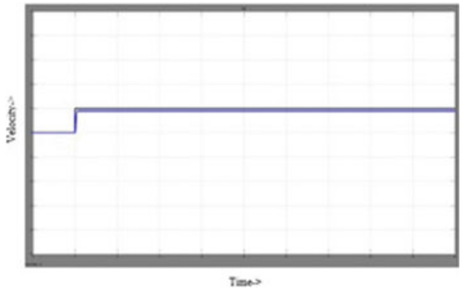
Closed loop with P Controller

A closed loop system with a proportional gain controller reduces the rise time, increases the overshoot, and reduces the steady-state error. If we use a greater proportional gain, the rise time and steady-state error will become even smaller (Figs. 11, 12, 13 and 14).

Closed loop with PI Controller

A closed loop system with a PI gain controller helps one achieve improvised damping with zero offset value resulting in no steady-state error (Figs. 15, 16, 17 and 18).

Fig. 11 Response of step input to the plant



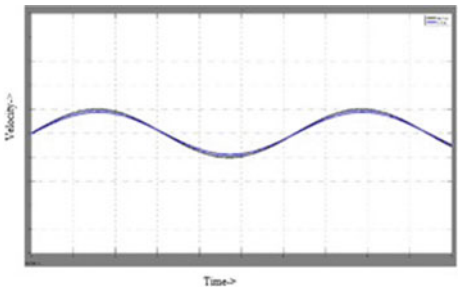


Fig. 12 Response of sinusoidal input to the plant

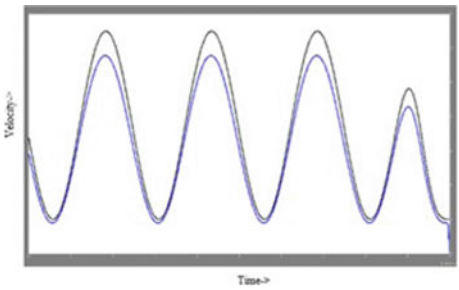


Fig. 13 Response of trajectory 1 input to the plant

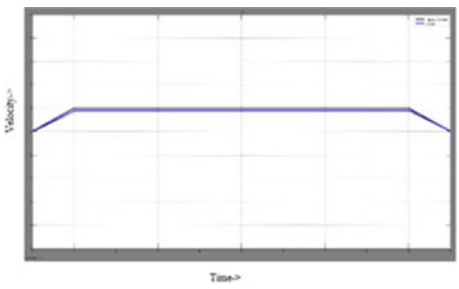


Fig. 14 Response of trajectory 2 input to the plant

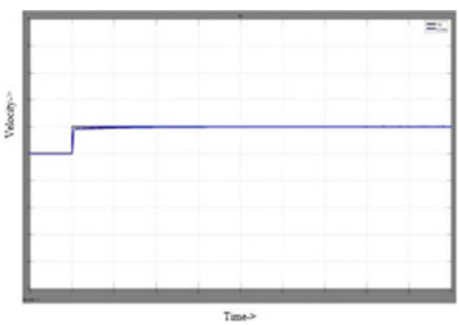


Fig. 15 Response of step input to the plant

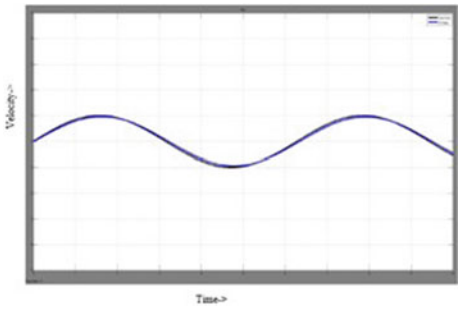


Fig. 16 Response of sinusoidal input to the plant

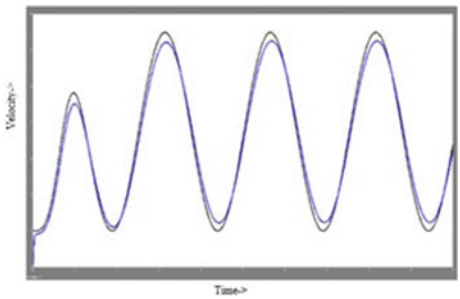


Fig. 17 Response of trajectory 1 input to the plant

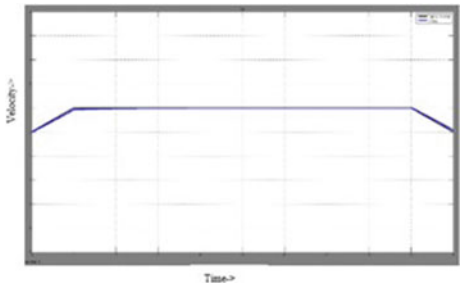


Fig. 18 Response of trajectory 2 input to the plant

Closed loop with PD Controller

A closed loop system with a PD gain controller improvises the transient state analysis factors such as decrease in maximum peak overshoot, rising time as well as settling time resulting in increased bandwidth (Figs. 19, 20, 21 and 22).

Closed loop with PID Controller

A closed loop system with a proportional-integral-derivative gain controller is preferred for implementation. The proportional gain controller helps decreasing the rise time. The integral gain controller results in elimination of steady-state error where the derivative controller reduces the maximum peak overshoot and settling

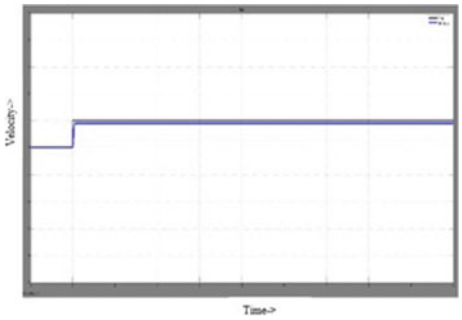


Fig. 19 Response of step input to the plant

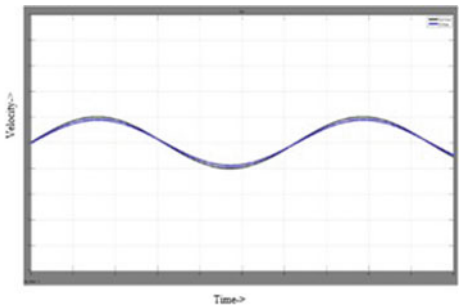


Fig. 20 Response of sinusoidal input to the plant

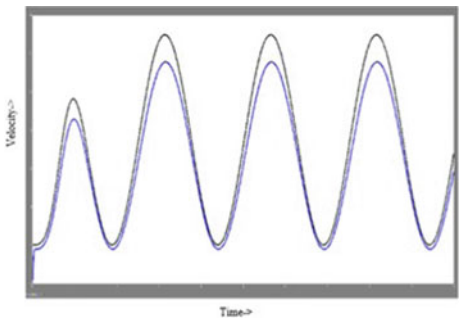


Fig. 21 Response of trajectory 1 input to the plant

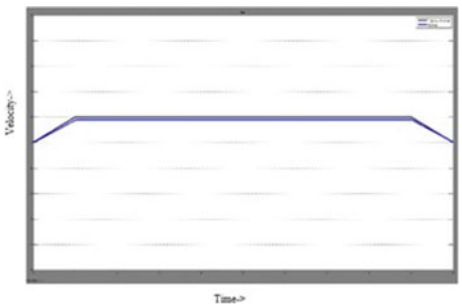


Fig. 22 Response of trajectory 2 input to the plant

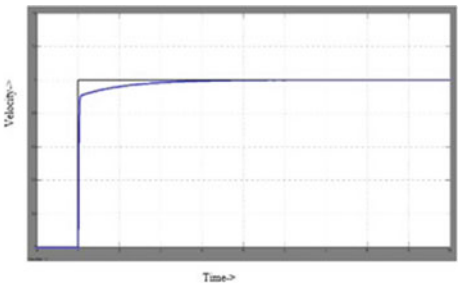


Fig. 23 Response of step input to the plant

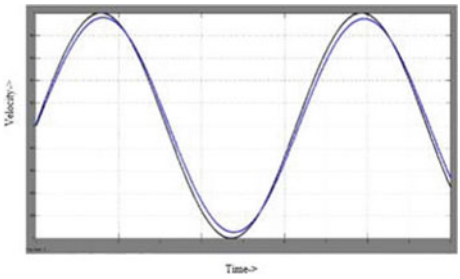


Fig. 24 Response of sinusoidal input to the plant

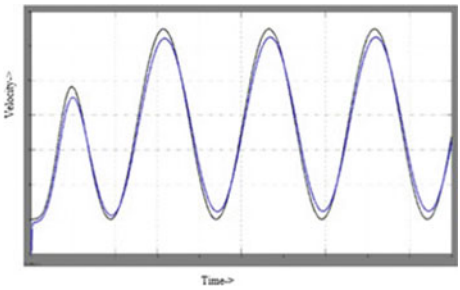


Fig. 25 Response of trajectory 1 input to the plant

time. Hence, altogether a PID controller implementation is executed (Figs. 23, 24, 25 and 26).

8 Conclusions

The paper is a complete representation of system identification, analysis, modeling, simulation, validation, and verification of speed-position control systems. Different trajectories are realized. During experimental analysis and simulation, it has been observed that due to generation of smooth trajectory, the synchronization error of the control system reduces. The analysis also demonstrates that the speed-position

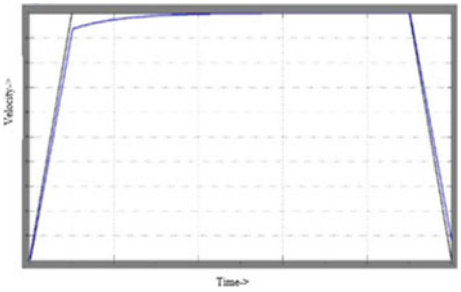


Fig. 26 Response of trajectory 2 input to the plant

control observed in the output gives improvised output as compared to the input. Experimental verification is done using software in loop simulation.

Model-based design has way more positive sides and the key advantage would be dealing with complex control systems in terms of efficiency; thus, at the same time, as a consequence, it limits the designer with a complex debugging if problem arise. It could be challenging for designers to deal with.

This system can be easily and efficiently be merged with Internet of things (IoT) for analysis. Considering model-based design as a key factor, putting up IoT along with embedded systems can implement a cost-effective smart intelligent system.

References

1. Jones, D., & Mckay, B. *Model-Based Design of Control Systems: Simulate and Test Before Committing to Hardware.* [Online]. Available: https://tinyurl.com/ybuyk4ge. Accessed April 12, 2020.
2. MathWorks. *Evolution of Model-Based Design in Aerospace.* [Online]. Available: https://tinyurl.com/y9jaorjk. Accessed May 8, 2020.
3. MathWorks. *Model-Based Design with Simulink*. [Online]. Available: https://tinyurl.com/ycnhbl66. Accessed April 15, 2020.
4. Mhase, V. (2012). *Integrated speed-position tracking with trajectory generation and synchronization for 2-axis DC motion control*. India: Pune.
5. Silva-Ortigoza, R., Hernandez-Guzman, V. M., Antonio-Cruz, M., & Munoz-Carrillo, D. (2015). DC/DC buck power converter as a smooth starter for a DC motor based on a hierarchical control. *IEEE Transactions on Power Electronics, 30*(2), 1076–1084.
6. MathWorks. *Introduction: PID Control Design.* [Online]. Available: https://tinyurl.com/ncdmkdt. Accessed April 13, 2020.
7. Ogata, K. (1997). *Modern control engineering* (3rd ed.). Tom Robbins: United States.
8. Sariyildiz, E., Yu, H., & Ohnishi, K. (2015). A practical tuning method for the robust PID controller with velocity feed-back. *Machines, 3,* 208–222. https://doi.org/10.3390/machines3030208.
9. Loucif, F. *DC Motor Speed Control using PID Controller*, June 2–5, KINTEX, Gyeonggi-Do, Korea.
10. Nagrath, I. J., & Gopal, M. (2003). *Control systems engineering* (5th ed.).

Design of Cardiovascular Disease Classification Using Wavelet Transform and Detection Using Neural Network

Neerajkumar S. Sathawane and Ulhaskumar M. Gokhale

Abstract In the previous few decades, maximum number of the studies has been centered on the process of medical specialty indicators. Routine clinical exercise creates huge quantity of biomedical signals like ECG in the course of observance of sufferers and for diagnostic functions [1]. Thus, automatic process systems are oftentimes utilized in clinical records analysis. Advanced methods will modify and advances volumes of facts. The medical doctor terribly oftentimes must check a patient's designation on the thought of analytical results in the course of examination. The coronary heart characteristic could also be analyzed through the electrocardiographic signal [2]. The popularity of ECG communication is predicated completely on its simplicity and non-invasiveness.

Keywords Electrocardiography (ECG) · Arrhythmia · Wavelet transform (WT) · Artificial neural network (ANN) · Multilayer perceptron (MLP)

1 Introduction

Electrocardiograms (ECGs) are electrical perspectives of the heart. The recording of associated electric signal picked up from the specific points from the body and presents the abnormal of action potentials of the coronary heart cells [1]. ECG comprises of waves P, Q, R, S and T. P-wave represents depolarization of atria, blood progress from atria to ventricles. Next PQ section is a progress of blood in the chambers atria to ventricles. QRS complex is an important part of ECG diagnosis and generated due to depolarization of ventricles. Repolarization of atria is not seen on the report; it is converged by way of QRS complex. T-wave is generated by the

N. S. Sathawane (✉)
G. H. Raisoni College of Engineering, Nagpur, MH, India
e-mail: niraj_51413@rediffmail.com

U. M. Gokhale
G.H. Raisoni College of Engineering and Technology, Nagpur, MH, India
e-mail: umgokhale@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_5

repolarization of ventricles. There is commonly glaring U-wave once the T-wave is over. U-wave is taken into account as overdue repolarization of cardiac muscles in ventricles and is every so often not seen [1].

2 What is a Disorder (CVD)?

ECG record consists of heartbeats that repeat sporadically. In every heartbeat, we are able to acknowledge many waves and inter-wave segments. The form and distance of those waves and link segments specify diseases like arrhythmias, heart muscle anemia and different heart diseases [3]. Existing techniques usually suffer for response to noise and incorrectness in access to new or ambiguous patterns. For medical follow, there is a requirement to evolve distinguisher that modifies nonlinear discrimination between categories, broken or irrespective inputs [4] (Fig. 1).

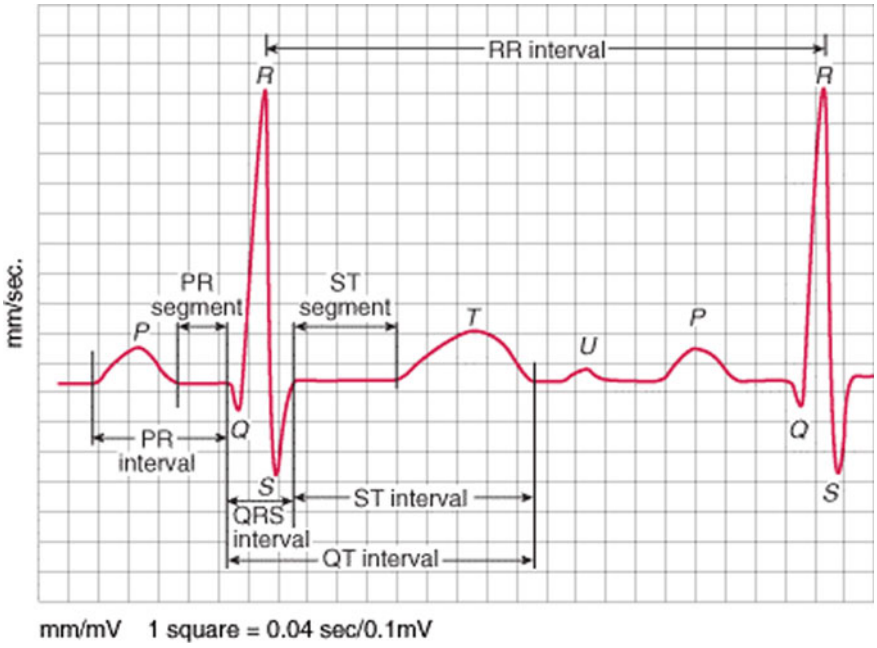


Fig. 1 Standard ECG wave

3 Wavelet Transform

The wavelet transform (WT) has emerged over recent years as a robust time-frequency analysis and signal cryptography tool favored for the interrogation of advanced non-stationary signals [5].

3.1 Continuous Wavelet Transform

Wavelet transform (WT) [1] is predicated as utilization and collection of competitive analyzing functions known as "wavelets". Wavelets break the signal into a coefficient of the input wave. Here, complex-valued function $\psi(t)$ denotes dilation and time shift. Every analyzing function $\psi\_(s, \tau)\ (t) = \psi((t - \tau)/s)$ can be explained in brief as a time localization drawn by shift τ also band outlined by a dilation factor *s*. Wave coefficients may be simply calculated and represented by the continual wavelet transform (CWT) [5].

$$\mathbf{CWT(s, \tau)} = \frac{1}{\sqrt{\mathbf{s}}} \int_{-\infty}^{+\infty} \boldsymbol{\psi}^* \left(\frac{\mathbf{t} - \boldsymbol{\tau}}{\mathbf{s}} \right) \mathbf{f(t)dt}$$

Here, input analyzing signal of an ECG is $f(t)$ also complex conjugate represented by *. Wavelet transform provides yield in the form of wavelet coefficients.

3.2 Mother Wavelets

The analytical use of this MW $\psi(t)$ depends on their characteristics linked with the application. These wavelets are selected because of their excellently matching shape and ease of using those in biomedical signals. CWT is part of wavelets that are expressly outlined by computational functions.

3.3 Discrete Wavelet Transform

The discrete wavelet transform (DWT) that could be a time-scale illustration of the numerical spikes that are obtained from digital filtering techniques is found to yield a quick computation of ripple rework [1]. The DWT that conjointly noted as decomposition by ripple filter banks is computed by serial low-pass filter (LPF) and high-pass filter (HPF) of the separate time domain signal shown graphically in Fig. 2 [5].

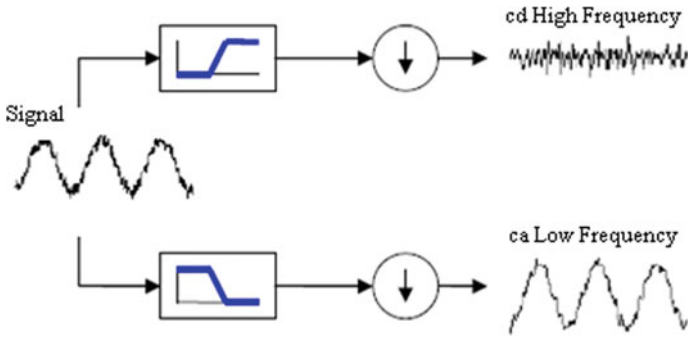


Fig. 2 Filter banks signal decomposition

3.4 Wavelets and Associated Families

With reference to one-dimensionality [6], and tend to identify the ripple ψ from the given function Φ, referred to as the scaling operator. ψ and Φ can be defined in brief as

- The whole of ψ is zero, $(\int \psi(x)\mathrm{d}x = 0)$ and is illustrating the details.
- The whole of Φ is 1, $(\int \Phi(x)\mathrm{d}x = 1)$ and is illustrating the approximations.
- The function ψ may be represented as $\psi(t)$ and Φ may be represented as $\Phi(t)$.
- $\hat{\psi}$ is Fourier transform of ψ, and $\hat{\Phi}$ is Fourier transform of Φ.

3.5 Features Extraction Using DWT

Features extraction could be a methodology of extracting and changing the computer file info into a collection of options known as feature vector, by reducing the information illustration pattern [7]. The options set to extract the relevant info from the computer file so as to perform the classification task. WT encompasses a varied window size, being broad at low frequencies and slim at high frequencies, so resulting in associate in nursing best time-frequency resolution all told frequency ranges.

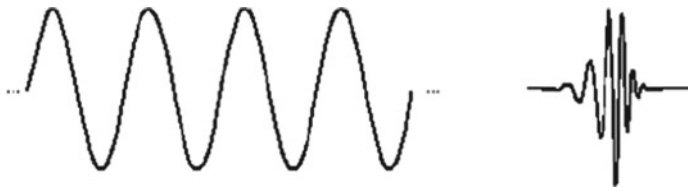


Fig. 3 Diagrams of sinusoidal signal and Daubechies wavelet [7]

From Fig. 3, the signals with intense variations can be easily computed compared to irregular moving ridge than with a swish sinusoid [6]. For a moving ridge of order N, the premise perform is often diagrammatical in the equation as

$$\psi(n) = \sum_{j=0}^{N-1} (-1)^j c_j (2n + j - N + 1)$$

4 Classification of Cardiovascular Diseases Using Ann

Automatic classification of ECG signal consists of different features of ECG in one cardiac cycle. The sluggish job for doctors is to check and interpret from long ECG records. So, to involve computers in this method is proposed for automated diagnosis of the abnormalities in the ECG. The prime motto of this method is attached to pattern recognition technique [8].

4.1 Classification of Cardiovascular Diseases Using ANN

Following parameters are set for training ANN model [9].

(1) Multi-layered feed-forward back-propagation NN is used for classification. (2) For fixing the number of hidden layers, ANN model with varied number of hidden layers are tried and empirically selected the model with two hidden layers as it gives the best performance. (3) Transfer function at hidden layers is "tansig", and transfer function at output layer is "logsig". (4) The network is trained with "trainlm" as it gives better results. (5) Network is trained with a select beat from different files of databases. (6) Classifier is tested for recognition and accuracy of classification of CVD. (7) Classification Epochs: 5000. (8) Goal: 0–0.00001. (9) Error function: mean square error (MSE). (10) Learning momentum (μ) = 0.001. (11) Validation fail-10. (12) Training dataset-60% (initial). (13) Test dataset-40% (remaining). (14) Validation dataset-60% (last with 20% overlap).

4.1.1 ANN Classifier Model (Features Extraction Using Decimation in Time Domain Approach and Wavelet Transform (WT) Domain Approach)

(1) Modular approach is employed for different features datasets (e.g., 21, 26, 32, 42 and 63) which are obtained after decimation.
(2) For each feature's dataset, different ANN models are designed and trained as depicted in Table 1.

Table 1 Different ANN models

Model	Input neurons	Neurons in hidden layers	Output neurons
M21	21	21	1
M26	26	26	1
M32	32	32	1
M42	42	42	1
M63	63	63	1

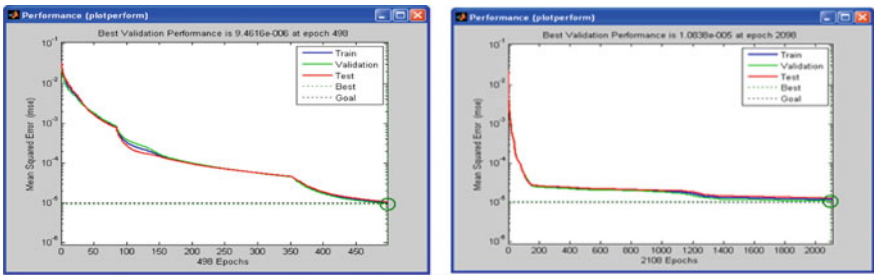


Fig. 4 Training performance for model **a** M32-features extraction using decimation. **b** Morphological and statistical features extraction

(3) Output = Maximum likelihood of (M21, M26, M32, M42, M63).

Five different models are prepared for 184 features and are trained. Training performance for model M21, M26, M32, M42 and M63 is recorded. Each model is trained thrice. Figure 4a shows performance plot of M32 model at the end of third training, and Fig. 4b shows morphological and statistical features extraction [10, 9].

4.1.2 Classification of Cardiovascular Diseases Using Adaptive Neuro-Fuzzy Inference System (ANFIS)

For training ANFIS classifier

(1) Subtractive clustering method is used to initialize FIS (genfis2). (2) Gaussian bell membership function is employed. (3) Cluster radius = 0.5. (4) Epochs: 1500.

Other parameters vary depending on feature extraction method.

ANFIS Classifier Model (Features Extraction Using Decimation in Time Domain Approach and Wavelet Transform (WT) Domain Approach)

Following information is used during training of ANFIS classifier ANFIS info:

Total nodes: 377, total linear parameters: 175.

Total nonlinear parameters: 336, total parameters: 511.

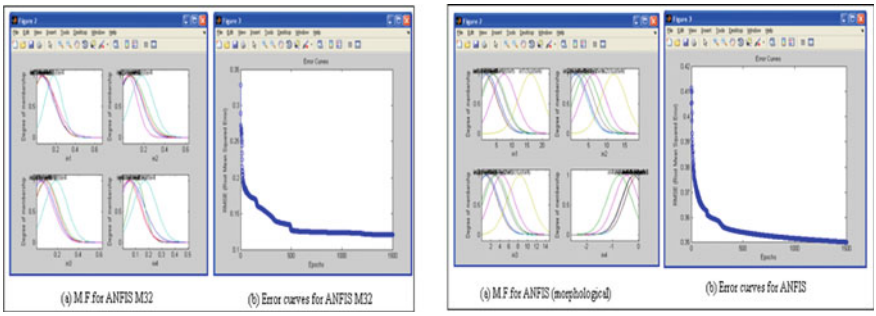


Fig. 5 **a** Membership function for ANFIS M32 and error curves for ANFIS M32, **b** Membership function for ANFIS (morphological) and error curves

Total training data pairs: 4140, total checking data pairs: 4140.

Total fuzzy rules: 7

Features extraction using decimation in time domain approach in Fig. 5a and WT in Fig. 5b.

5 Results

In this research, Lead-II ECG signals are obtained from databases for performance evaluation and training of the ECG classifier. The signal window length (dataset) is arbitrarily chosen to be 1 s. Arrhythmias to be classified are NCR, EH, APB, FHR, V-Fib and PSVT. MIT-BIH Arrhythmia Database used to evaluate our algorithm (Table 2).

Table 2 Datasets for six classes

ECG Signals	Class	Training set	Testing set	Total
Normal cardiac rhythm (NCR)	1	2800	3000	5800
Ectopic heartbeat (EH)	2	550	600	1150
Atrial premature beats (APB)	3	550	600	1150
Ventricular fibrillation (V-Fib)	4	80	100	180
Fast heart rhythm (FHR)	5	80	100	180
Paroxysmal supra ventricular tachycardia (PSVT)	6	80	100	180
Total		4140	4500	8640

Table 3 Testing results for ANN-decimation model

ANN-decimation model									
Classes	T. Set	NSR	PVC	APC	VF	VT	SVT	Unk	Computation
NCR	3000	2976	0	20	0	0	0	4	149 time (s)
EH	600	0	590	0	0	0	4	6	
APB	600	11	0	589	0	0	0	0	
V-Fib	100	0	0	0	96	0	0	4	
FHR	100	0	0	0	4	94	0	2	
PSVT	100	0	5	0	0	0	95	0	

Table 4 Testing results for ANN-morphological model

ANN-morphological model									
Classes	T. Set	NSR	PVC	APC	VF	VT	SVT	Unk	Computation
NCR	3000	2892	0	92	0	0	0	16	169 time (s)
EH	600	0	568	0	0	20	0	12	
APB	600	18	0	572	0	0	10	0	
V-Fib	100	0	0	0	88	8	0	4	
FHR	100	0	4	0	7	87	0	2	
PSVT	100	0	7	0	0	0	89	4	

5.1 Testing Results

Tables 3 and 4 show the testing results for ANN-decimation model.

A total of 8640 datasets for six classes, out of which 4140 datasets were used for training and 4500 datasets used for testing, are shown in Table 2. Apart from this, some samples from local hospital are also collected [11].

5.2 Performance of the Classification

The benefit of ECG beat classification is to evaluate the performance of a method by appropriation. Here, we have focused on three statistical indices Sensitivity (Se), Specificity (Sp) and Accuracy (A). Tables 5 and 6 show performance of the classification for ANN-decimation model, and Table 7 shows the comparison between different methods based on number of beat types and overall accuracy (%).

Table 5 Performance of the classification for ANN-decimation model

Performance: ANN-decimation model						
Parameters	NCR	EH	APB	V-Fib	FHR	PSVT
Sensitivity in %	99.33	99.33	98.17	100	95.92	95
Specificity in %	99.63	99.16	96.72	96	100	95.96
Accuracy in %	99.2	98.33	98.17	96	94	95

Table 6 Performance of the classification for ANN-morphological model

Performance: ANN-morphological model						
Parameters	NCR	EH	APB	V-Fib	FHR	PSVT
Sensitivity in %	96.92	96.60	95.33	91.67	88.77	92.71
Specificity in %	99.38	98.10	86.14	92.63	75.65	89.9
Accuracy in %	96.4	94.67	95.33	88	87	89

Table 7 Comparison between different methods

Reference	Method	Number of beat types	Overall accuracy (%)
Proposed method 1	ANN-decimation	6	98.67
Proposed method 2	ANN-morphology	6	95.47
Proposed method 3	ANFIS-decimation	6	95.47
Proposed method 4	ANFIS-morphology	6	96
[3]	ANFIS-DWT	8	94
[12]	Fuzzy-DWT	8	93.13

5.3 Calculation of Accuracy, Sensitivity and Specificity

Sensitivity associated with the test's capability to handle positive results. Specificity associated with the capability of the test to handle negative results.

$$\text{Sensitivity } \% = \frac{\text{Complete Tested Sample set} - \text{FN}}{\text{Complete Tested Samples Set}} \times 100$$

$$\text{Specificity } \% = \frac{\text{Complete Tested Sample set} - \text{FP}}{\text{Complete Tested Samples Set}} \times 100$$

$$\text{Accuracy } \% = \frac{\text{Complete Tested Sample set} - \text{FP} - \text{FN}}{\text{Complete Tested Samples Set}} \times 100$$

Here, FP denotes for false positive and FN denotes for false negative.

5.4 Comparison Between Different Methods

See Table 7.

6 Conclusions

Various electrocardiogram analysis ways supported discrete wavelet transform (DWT), and ANN was analyzed and balanced and outperforming unit evolved which displays excellent and foolproof results than conventional printed results. Out of the all studied models, ANN-reduced primitive model is fabulous and computationally quick. The system could also be used for knowledge analysis in ambulatory electrocardiography device that unceasingly records the heart rhythms for twenty-four–forty-eight hours. Also we will share this conclusion over mobile with consultant.

References

1. Sarma, P., Nirmala, S. R., & Sarma, K. K. (2014). ECG classification using wavelet subband energy based features. In *2014 ICSPIN*.
2. Djahida, M., Talha, K., & Slimmane, S. (2018). Neural networks and SVM for heartbeat classification. *IEEE Sensors Journal*, pp. 277–286.
3. Ranaware, P., & Deshpande, R. (2016). Detection of arrhythmia based on discrete WT using ANN and SVM. In *ICCSP*.
4. Pal, S., & Mitra, M. (2010). *Detection of ECG characteristic points using multiresolution WA based SCM* (pp. 255–261).
5. Rai, H., & Trivedi, A. (2012). De-noising of ECG waveforms using MWT. *IJCA, 45*(18), 55–60.
6. Vaneghi, M., Oladazimi, M., Shiman, F., Kordi, A., Safari, M., & Ibrahim, F. (2012). A comparative approach to ECG FEM. In *2012 ICISMS* (pp. 252–256).
7. Karpagachelvi, S., Arthanari, M., & Sivakumar, M. (2010). ECG feature extraction techniques-a survey approach. *IJCSIS, 8,* 76–80.
8. Kaur, M., & Kaur, G. (2015). Adaptive wavelet thresholding for noise reduction in electrocardiogram (ECG) signals. *IJCSNS, 3,* 138–143.
9. Vidya, K. S., Acharya, U. R., Ghista, D. N., & Lim, W. J. E. (2015). Computer-aided diagnosis of diabetic subjects by HRVS using DWT method. *Knowledge-based systems, 81,* 56–64.
10. Abo-Zahhad, M. (2011). ECG signal compression using discrete wavelet transform. In *Discrete wavelet transforms-theory and application* (pp. 143–168).
11. Massachusetts Institute of Technology, MIT-BIH arrhythmia database.
12. Liu, S. H., Chang, K. M., & Wang, J. J. (2011). *An automatic system for ECG arrhythmias classification* (pp. 2290–2294). IEEE.

Enhanced Lightweight Model for Detection of Phishing URL Using Machine Learning

Siddhesh Masurkar and Vipul Dalal

Abstract In the past few years, many forged sites are created on the Internet to impersonate genuine websites, with the goal of acquiring sensitive and valuable information from the people and firms. Such type of attack which is carried online is known as phishing, and it continues to carry serious threats for consumers and business and the many shareholders' hundreds of million dollars. Many cyber intrusions are successful through phishing. Phishing attacks are the attacks wherein Internet users are fooled by disguising a fake website as a legitimate website. As technology progresses, the phishing detection methods need to get advanced and there is a terrible requirement for improved mechanism to avoid, check, as well as detect these phishing attacks. Thus, useful remedial and preventive measures that can precisely find out phishing sites are a machine learning approach. Machine learning (ML) is a current tool for data analysis and lately has revealed capable leads in fighting phishing problems. Although many varieties of classification algorithms for detecting phishing are proposed, examined, and analyzed in many papers, it observed that advanced and depth of phishing threat are being continuously increased at steady rate. In this paper, we project a new method called enhanced lightweight model for phishing site detection. We further suggest the use of various URL features and boosting algorithm. We will analyze and examine the impact of boosting algorithm against the performance of other classification algorithms to find better result.

Keywords Phishing websites · Features · Boosting · Machine learning algorithm · URL · Attack · Classification

S. Masurkar (✉) · V. Dalal
Vidyalankar Institute of Technology, Mumbai, India
e-mail: siddhesh1199@gmail.com

V. Dalal
e-mail: vipul.dalal@vit.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_6

1 Introduction

In today's world, the Internet plays significant role in everyone's life. It has its convenience in respect of all daily routines like entertainment, communication, bank transaction, online shopping, etc. It also has its advantages and disadvantages. Every device is connected to the Internet so as to control it from remote place. Therefore, in contemporary society, keeping personal information safe on such device or computer is more difficult for every person. Any person with bad intent and sufficient knowledge can gain the access to such device or computer or to spread virus or malware in any system causing.

There are many crimes that happens in this virtual world. The main reason is that the attacker can be anonymous behind the curtain of this virtual world while carrying such bad activities. Phishing is a kind of attack wherein focus is on sensitive information and data entered by the user on a site which is designed in such a way that it disguises user as it is a trusted and real site which he used to surf. Once such information is obtained, it can be misused to make happen scams and make transactions. The APWG phishing Activities Trends Report says that in last semester, approx. 4 million phishing sites are reported I tear 2019. Most of the attacks were on SaaS and webmail which is 30.8%, payment site (19.8%) and financial institution (19.04%). Figure 1 shows the steady increment in numbers of phishing sites detected in THE period of 2017 to 2019 [1].

Many studies have been proposed to curb the problems caused by phishers. However, there is still not a promising way to find the phishing sites. Phishing detection techniques mainly have two main ways to find fake sites: heuristic-based approach and blacklist-based technique. This paper provides a novel technique for detecting phishing sites based on boosting algorithm which aids in determining the given site is a genuine site or not.

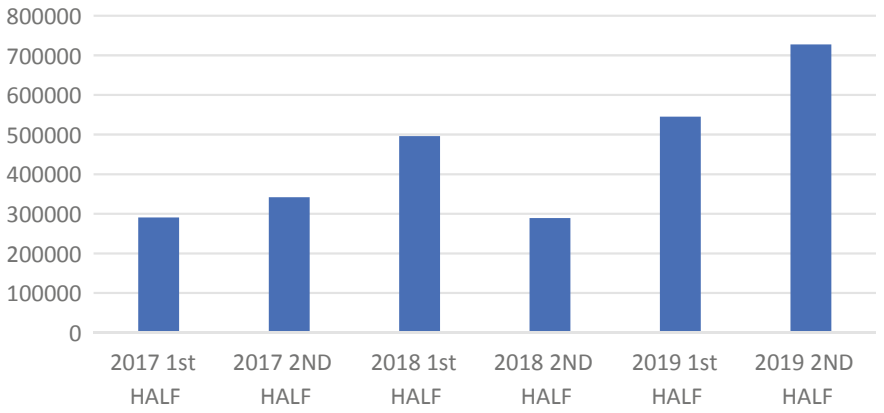


Fig. 1 Numbers of phishing sites detected

One method is to maintain a blacklist of phishing sites. But the problem with this approach is that this list of malicious website is not perfect because new malicious URLs are made with some changes continuously and progressively. Thus, such methods are expected that they should categorize a fresh, never before seen URL as a phishing web page or a real one. Such results are normally using machine learning knowledge approaches in which a system is developed to classify any given phishing site passed through a model built using large parameter of well-known attacks usually known as training sets.

The proposed system extracts parameters of the URL from the URL given by the programmer. The extracted features are given to boosting algorithm. Those features will be used by the boosting algorithm to find out a site requested by programmer which is a phishing site or real one.

One of the main issues found in available literature with building a solution using machine learning for this problem is that in public domain, there are very limited training datasets for available phishing URLs. Because of this, for dataset which is currently available, studies are required to assess the impact and effectiveness of machine learning approaches. This need is fulfilled by this proposed work. Definitely, the primary aim of this system is to match and analyze the outcomes of the widely used algorithms on the same phishing dataset which is created by processing large numbers of URL. In this system, large dataset is obtained in which features have been extracted from the URLs data, and those parameters have values based on some conditions. Widely used machine learning algorithms are going to be verified for the intent of dividing URLs such as random forest, Naive Bayes classifier, and SVM.

In this approach, phishing website is found, and it helps in decreasing the loss and damage caused by such malicious attacks. The paper is divided into sections as per the following: Sect. 2 shows the literature review regarding the topic. The problem statement and objective of the proposed system are explained in Sect. 3. Section 4 discusses the architecture of the proposed system. The method of assessment is explained in Sect. 5 followed by conclusion in Sect. 6.

2 Literature Review

Mao et al. [2] used different method in which CSS properties of web pages have been extracted from the URL since every page has CSS content. They have used SVM, DT, AB, and RF classification model to train 19,918 phishing sites and 4133 benign sites. They have obtained best accuracy of 95% with RF and SVM methods.

Zhang et al. [3] show that the authors have extracted 11 features from URL and web and used ELM technique. It has two stages. They extracted 11 features from five classes and four different classifiers, namely AdaBoost, random forest, bagging, and SMO which were used. First is building content classifier, and second is combining text and other hybrid classifiers, and it has third party dependency to extract the features. The limitation associated with their approach is they lack real-time demand of phishing detection.

El-Alfy [4] projected the multi-forward feedforward ANNs and their variants to detect the phishing sites. They have used 30 features which are extracted from the URLs. Those features are categorized into four classes and based on each class along with feature selection technique. It has been further observed that when all features are used for detection, it is providing best result among all which is 96.07%. But drawback of this system is that it consumes more time to find URL which is phishing or not.

Parekh et al. [5] projected URL identification strategy with the help of Rstudio tool to provide better analysis. They have obtained dataset form phishtank.com and are using only limited dataset. Out of 31 features, they are using only eight features. On these eight features, information gain and ranker method are used to found important features. Dataset is divided on 70:30 ratio for training and testing, respectively. The random forest classifier is used in the approach which is providing accuracy of 95%.

MacHado et al. [6] explain the system to detect the phishing URL using C4.5 decision tree algorithm. The build system is robust and precise, and URL features are making it more accurate. They extracted eight features and built a model which applied on three datasets. They attained average accuracy of 89.4%.

Aburrous et al. [7] in this paper used 27 features to build the classifier which is based on the fuzzy system. Their approaches mainly concentrate on the banking domain. They have used CBA and MCAR classifiers.

The authors in [8] projected email classification model with mining concept which extracts 23 keywords from the email. They have used 2500 emails out of which 1260 emails and others are valid emails. In text mining, they have mined some keyword, and using *t*-statistic methods, the statistic values of that keyword are calculated. Using those features, model is built using seven classifiers, and they all evaluated on accuracy, specificity, and sensitivity with and without feature section. It is observed that all models using all features are giving best accuracy of 98.12%. The genetic programming is yielding best result.

The study [9] presents the use of Bayesian classifier to detect phishing email. They are removing HTML tag and stop word. After that, they are tokenizing the URL and calculating the token frequency in URL. Accordingly, Bayesian model is built and evaluated. That model is tested on three datasets with highest accuracy which is 96.46%. They used about 4700 emails as a dataset.

Form et al. in [10] used nine features which consist of 5 email structure and 4 is based on sender behavior-based features. They have considered 500 phishing and 500 valid emails, and SVM is used to build the model along with support vector machine classifier to determine emails as genuine or not. The accomplished accuracy of model is 95.25% in results, but its drawback is that comparatively, very small training dataset (1000 emails with 50% spam and 50% ham) is used.

Peng et al. [11] present an approach to distinguish attacks of phishing email using NLP and classifiers. The text semantic analysis is used to find harmful content of the email. They have used 5009 phishing emails and 5000 valid emails.

A method of detection using heuristic features by authors [12] used only ranking features which are obtained from the URL, namely PageRank, AlexaRank, and AlexReputation along with some basic features like primary domain, subdomain,

and path domain. They are calculated six values of heuristic. They are using dataset of 11,660 phishing sites. Then they have calculated weights of these heuristics. The flow of the approach is divided into six phases which are implemented in PHP and MYSQL. Ten datasets are tested against the build model. According to result, the threshold value of 0.4, 0.5 and 0.6 is having accuracy above 97% which is 97.16%, 97.10%, and 97.09%, respectively.

3 Problem Statement

As per the report on anti-phishing working from July to December 2019, 254,912 different phishing attacks were observed globally in the second half of 2019 [13, 14]. These attacks targeted industries mostly like bank, e-commerce, and money transfer companies and websites. Seventy-five percent of phishing websites use five top-level domains in same period which comprises .com, .uk, .ru, .xyz, and .net.

The most exploited industry sector during the period of July to December of 2019 is SaaS/webmail was overtaking the financial and banking and other service sectors. The attackers targeted customers because of customer's plastic card valuable data and other information in respect of the personal bank accounts. The spam emails can be sent by the attacker from the hacked user's account. The business email compromise (BEC) scam is one of the major problems of 2019 [14]. In this scam, industries are fooled by the attacker into removing large amounts of money using a technique called as spear phishing techniques. According to RSA Fraud Report, phishing accounted for 37% of all fraud attacks in Quarter 4 of 2019. Phishing volume has increased 6% since last year.

In literatures, authors have used many methods and achieved the accuracy in the range of 86–98% with varying sets of features ranging from 4 to 31 with feature selection or not as shown in Table 1. Inclusion of large number of features may take large amount of time to determine which URL could increase the complexity and shows the dependency on third party. Since life cycle of phishing site is small, there is a need of fast and intelligent detection mechanism.

3.1 Objectives

The fundamental objective of this system is to develop a phishing detection mechanism. The system can work on limited features to provide with or without feature selection to provide better accuracy. Phishing detection mechanism intends to use information mining on URL dataset to aid the distinguish phishing URLs.

Table 1 Comparison between existing works

Existing work	Number of features	Approach used	Dataset used	Accuracy (%)
Mao et al. [2]	36	SVM, AB, RF, DT	19,918 phishing 4133 benign	95
Zhang et al. [3]	11	AdaBoost, bagging, random forest, and SMO	7788 URLs	98
El-Alfy [4]	30	PNNs and K-medoids clustering	4898 benign 6157 phishing	96.07
Parekh et al. [5]	31	Random forest	–	95
MacHado et al. [6]	8	C4.5 algorithm	3000 phishing and legitimate	89.4
Aburrous et al. [7]	27	C4.5, JRip, PART, PRISM, CBA, and MCAR	1006 phishing	86.3
Pandey et al. [8]	23	DT, LR, SVM, MLP, GMDH, PNN, and GP	1260 phish 1240 valid	98.12
Rathod et al. [9]	–	Bayesian classifier	4700 dataset	96.46
Form et al. in [10]	9	C4.5 algo	500 spam 500 valid	95.25
Peng et al. [11]	9	Seahoud and Netcraft	5009 phishing 5000 valid	95

3.2 Specific Objectives

- Extract the URL features from the dataset and select best features.
- Build the model using boosting algorithm and other widely used classifiers.
- Evaluate the performance of boosting classifiers on different parameters against traditional classifiers.

4 Proposed System

The paper's main idea is to project a detection system for phishing sites which can identify phishing websites from URL using features set of ten features which include lexical features of the URL only, and there is no involvement of any services from third party. This method has following main blocks, namely (1) feature extractor which is accountable for removing the appropriate set of lexical feature and (2) a model that has a role of detection of phishing site by testing test set. Each component is described in details as follows.

4.1 Structure of the URL

URL stands Uniform Resource Locator. A URL is an address given to unique resources on the web. It is human-understandable text that was used to replace the IP addresses that computers understand and use to communicate with other systems in communication. A URL comprises of a protocol, domain name, and path (which contains the specific subfolder structure where a page is placed) and has the following basic format (Fig. 2):

Protocol: The protocol is a format in which computer browser exchanges data with the server while transmitting or requesting a page or content of the pages. Nowadays, widely used protocol is HTTP which means Hypertext Transfer Protocol. Alternative to HTTP is commonly used protocol HTTPs which stands for Hypertext Transfer Protocol Secure. It means your web browser can provide security by encrypting any information, data that is provided to browser so that phisher cannot understand that who try to take, seize the page during transfer of data.

Domain: A domain name is an inimitable part which recognizes a site, pages on the Internet, for example, londonbridge.co.uk. A top-level domain (TLD) is contained in domain name.

Subdomain: The main domain has a small part which is known as a subdomain. For instance 'indiatimes.timesofindia.com' indiatimes is subdomains of the domain name timesofindia.com.

Port: The URL has a port number which is seldom observable, but every time it is required. The position of port is always after TLD, which is distinguished by a colon (:). When there is port mentioned in the URL like in most of the cases, then by default HTTP is protocol and port 80 is used.

Path: The path normally denotes to a file or directory which can be accessed from the web server, e.g., /directory/sample/data.php. seldom the name of the file which is not specified in the URL.

Parameter: It is pieces of data which exist in URL in the form of query string. As per the query searched on the Google, the parameters start after the special character which is usually question mark and are divided by an ampersand (&); therefore, it can be interpreted separately and used to show body of the page.

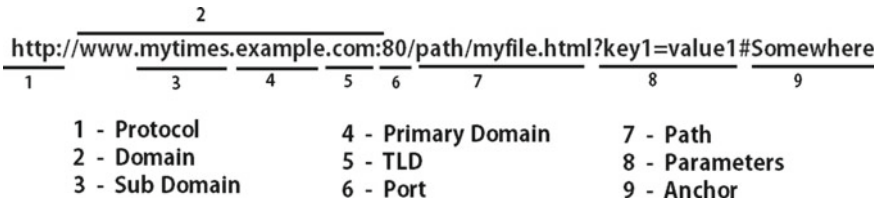


Fig. 2 Structure of the URL

Anchor: An anchor is called as an internal page reference, which provides some information that can be processed. Any URL is ended with an anchor, and it starts with a hash (#) character following an identifier. It states to a part of data *within* a page.

4.2 URL Lexical Features

In our proposed system, we are using limited sets of URL features from 23 obtained features. Different parts of the URL give valuable information using which spoofed URL can be created quickly. Therefore, the information obtained from the URL can be used to identify whether URL is harmful or real. For this purpose, 23 different features are extracted from the URL. The 23 features which are extracted from the URL are majority lexical features except rank country and rank host. Those URL features are first separated by tokenization and parsing with the help of special character present in the URL. Further by processing it, values corresponding to those features are obtained. Those features are as follow rank host (f1), rank country (f2), length of URL (f3), length of host (f4), numbers of dots (f5), executable file in URL (f6), presence of '@' in URL (f7), presence of '/' in URL (f8), presence of '–' in URL (f9), phishing TLD in URL (f10), presence of HTTP in path (f11), shortening services used (f12), average length of the token (f13), numbers of token in URL (f14), largest token size (f15), average length of domain token (f16), domain token count (f17), largest domain (f18), average length of the path (f19), path token count (f20), largest path token (f21), presence of security sensitive words (f22), and presence of IP address (f23). Among these features, features f5, f6, f7, f8, f9, f10, f11, f12, f22, and f23 are having value either 0 or 1. Other features are having value greater than 0.

4.3 Flow of the Model

The proposed model is shown in Fig. 3. It uses the URL features which are listed here above to determine phishing site.

The proposed system has two phases: pre-processing phase and detection and evaluation phase. The dataset of URLs which include valid as well as phishing URLs is given as input. These URLs are first inputted to the feature extractor. Above-listed features f1 to f3 are extracted after tokenization and processing of URL. The values of these features could be one of 1 or 0 depending on the certain conditions, and features f1, f2, f3, f4, f13, f14, f15, f16, f17, f18, f19, f20, and f21 have value equal to and greater than 1. These features and its associated values are saved in a .csv file. Accordingly, feature dataset is created. Using feature dataset, model is trained by using boosting algorithm. Similarly, other models are trained using Naive

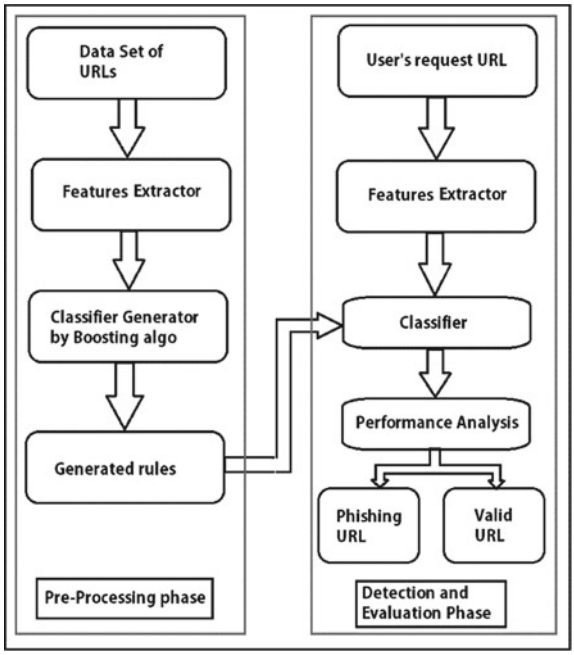


Fig. 3 Proposed model data flow

Bays, SVM, and random forest methods. These models will also be saved in .pkl for performance evaluation of the boosting algorithm.

In detection phase, URL which is to be tested is given as input. Again using same conditions, features are extracted and values are assigned after processing. Those values are given as an input to the model of various classifiers. Using classifier model, whether given URL is fake or valid is classified using boosting algorithm. If given URL is determined as phishing URL, then alert is made to the user about the classification result.

In evaluation phase, separate lists of URLs stored in a file like test_dataset1 and test_dataset2 are given to the feature extractor which subsequently extracts the features and value are assigned to those features. Extracted feature and its corresponding value are stored in .csv file. Using these test dataset, the accuracy of the saved models can be tested and its performance is compared on various parameters. In this system, random forest, SVM, Naïve Bays, and boosting algorithm are chosen to test and compare the phishing detection model.

5 Evaluation

Datasets of phishing sites and legitimate sites are collected from publically available sources [15–17] from the year 2018 to 2019. We have decided to use support vector

Table 2 Confusion matrix

	Predicted positive class	Predicted negative class
Actual positive class	**TP** (True Positive)	**FN** (False Negative)
Actual negative class	**FP** (False Positive)	**TN** (True Negative)

machine (SVM), random forest, and Naive Bays along with boosting algorithm. These are the simplest and most widely used models because of their effectiveness and good performance in binary classification.

To evaluate and compare the performance of model of boosting algorithm with other models, the following evaluation matrix is used, accuracy, precision, recall, and f-score. Accuracy measures the ratio of websites which are correctly predicted. Precision measures the fraction of websites correctly predicted as phishing. Recall metric measures the fraction of phishing websites identified by the model.

Table 2 shows the confusion matrix in which True Positive (TP) is a case in which a classification model correctly finds a website as phishing and True Negative (TN) is a case where a website is wrongly classified as valid URL. False Positive (FP) is a case where a website is wrongly classified as phishing, and lastly False Negative (FN) is when the model wrongly classified a website as valid URL while it is actually phishing. The mathematical equations of the performance metrics are given below, respectively.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{3}$$

$$\text{F - measure} = \frac{2 * \text{Presicion} * \text{Recall}}{(\text{Presicion} + \text{Recall})} \tag{4}$$

6 Conclusion

The most significant way to guard from the stealing confidential data and other information is the user awareness. The attacker is making phishing site in such a way that it really looks like a legitimate site and user is disguised to enter its private and sensitive information.

In this paper, we propose to detect phishing website using boosting algorithm. In our approach, it relies on analysis of the URL and its processing which will give you lexical features. In this method, we will be only dependent on URL or lexical features which are very easy to acquire and processing is simple. As a result, boosting algorithm could be more effective and accurate in finding phishing sites.

The present method is not using other features like HTML page features, event-based features, and domain features in phishing URL detection. Though these features are also important group to be considered, out future work is based on incorporation of those remaining features.

References

1. APWG report available at: https://docs.apwg.org/reports/apwg_trends_report_q3_2019.pdf. Accessed April 24, 2020.
2. Mao, J., Bian, J., Tian, W., et al. (2019). Phishing page detection via learning classifiers from page layout feature. *Journal on Wireless Communications and Networking,* 43.
3. Zhang, X., Zeng, Y., Jin, X., Yan, Z., & Geng, G. (2017). *Boosting the phishing detection performance by semantic analysis.*
4. El-Alfy, E. S. M. (2017). Detection of phishing websites based on probabilistic neural networks and K-Medoids clustering. *Computer Journal*. Peng, T., Harris, I., & Sawa, Y. (2018). Detecting phishing attacks using natural language processing and machine learning. In *Proceedings of 12th IEEE International Conference semantic computing ICSC 2018*(Vol. 2018, pp. 300–301).
5. Parekh, S., Parikh, D., Kotak, S., & Sankhe, P. S. (2018). A new method for detection of phishing websites: URL detection. In *Second International Conference on Inventive Communication and Computational Technologies (ICICCT)* (pp. 949–952)
6. MacHado, L., & Gadge, J. (2017). Phishing sites detection based on C4.5 decision tree algorithm. In *2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017* (pp. 1–5).
7. M. Aburrous, M. A. Hossain, K. Dahal, and F. Thabtah, (2010). Predicting phishing websites using classification mining techniques with experimental case studies. In *Proceedings of 7th International Conference Information Technology* (pp. 176–181).
8. Pandey, M., & Ravi, V. (2012). Detecting phishing e-mails using text and data mining. In *2012 IEEE International Conference on Computational Intelligence and Computing Research*
9. Rathod, S. B., & Pattewar, T. M. (2015). Content based spam detection in email using Bayesian classifier. In *IEEE ICCSP Conference.*
10. Form, L. M., Chiew, K. L., & Tiong, W. K. (2015). Phishing email detection technique by using hybrid features. In *2015 9th International Conference on IT in Asia (CITA)*
11. Peng, T., Harris, I., & Sawa, Y. (2018). Detecting phishing attacks using natural language processing and machine learning. In *Proceedings of 12th IEEE International Conference Semantic Computing ICSC 2018* (pp. 300–301).
12. Nguyen, L. A. T., To, B. L., Nguyen, H. K., & Nguyen, M. H. (2014). A novel approach for phishing detection using URL-based heuristic. In *2014 International Conference on Computing, Management and Telecommunications (ComManTel)* (pp. 298–303).
13. PhishTank. Online available at: https://www.phishtank.com/
14. APWG report available at: https://docs.apwg.org/reports/apwg_trends_report_q4_2019.pdf. Accessed April 24, 2020

15. Parekh, S., Kotak, S., Parekh, D., & Sankhe, S. (2018). A new method for detection of phishing websites: URL detection. In *Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018)*. IEEE Xplore Compliant—Part Number: CFP18BAC-ART; ISBN: 978-1-5386-1974-2
16. The Web Information Company. Online available at: www.alexa.com
17. DMOZ Open Directory Project. Online available at: http://www.dmoz.org

Detection of Spastic Cerebral Palsy Using Different Techniques in Infants

Nancy Sukhadia and Pariza Kamboj

Abstract Cerebral Palsy (CP) is a complicated neurological condition in infants. CP refers to brain injuries that affect different body parts. It must be detected in early childhood as early detection can help to improve outcomes through therapy-based interventions. Generally, pediatricians detect CP based on the physical movements of infants, but sometimes, it becomes difficult to detect this way at the early stages. The absence of "fidgety movement" is a strong predictor to detect CP using the combination of video and sensors. There are four kinds of cerebral palsy in which spastic cerebral palsy is very common in infants. It refers to muscles-related problems like muscle stiffness and muscle weakness. Spastic cerebral palsy is detected by measuring the angle between joints or other parameters like stride length, leg length, age, and cadence to name a few. Spastic cerebral palsy is typically detected using surface electromyography (sEMG), using sensors, namely accelerometer, inertial measurement unit (IMU), or using machine learning techniques. sEMG can be used only for recording the electrical activity of muscle tissues or visual representation. Impulsive extreme movements of lower and upper limbs are being observed using accelerometer, but it is not helpful to get postural information. Wearable IMU sensor consists of three sensors, namely accelerometer, gyroscope, and magnetometer. IMU records movement facts in three dimensions at numerous samples per second and forms any continuous hours without any mediation. This paper aims to study various conventional and modern techniques to examine CP and compares them based on various parameters. After a thorough study of various methods, it was found that IMU sensors provide the best way to detect CP in the case of infants as wireless IMU is a low cost, non-invasive, and easy to wear solution. They also provide 360° measurements through three sensors each on three axes which provides more accuracy in comparison with any other method.

N. Sukhadia (✉) · P. Kamboj
Sarvajanik College of Engineering and Technology, Surat, India
e-mail: nancysukhadia27@gmail.com

P. Kamboj
e-mail: parizakamboj@scet.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_7

Keywords Infants · Cerebral palsy · Joint angel · Spasticity · sEMG · Accelerometer · IMU

1 Introduction

The number of preterm children born in many industrialized countries has been increased tremendously. In India, 3,341,000 babies are born too soon each 12 months and 361,600 children underneath 5 years of age die due to direct premature complications [1]. Early infancy is an essential phase of an infant's existence, as many important major developmental mileposts are achieved progressively in the first six months which lay down the foundations for their imminent life [2]. However, these preterm infants are an expanded danger of growing illnesses like motor dysfunction, retinopathy, and apnea to say a few. Cerebral palsy (CP) is a prevalent motor dysfunction in preterm infants which is found in 4–20% of premature children based on their gestational age [3]. Neurodevelopment assessment in an infant is important to detect diseases at an early stage to save their life.

The word cerebral palsy is made up of two words cerebral and palsy. The initial one refers to the cerebrum, the uppermost area of the brain, and later one refers to the resultant disorder because of brain injury, which is shown in Fig. 1. The term "cerebral palsy" also refers as CP is a group of permanent cerebro-motor dysfunctions that appears in early infancy [4].

Cerebral palsy affects body movements and coordination permanently, but it is not degenerative. CP does not get worse over time, though the exact symptoms can change over a human's lifetime. Signs and symptoms of CP also vary among infants and overtime in cerebral palsy. Symptoms of cerebral palsy include poor coordination, oral problem, stiff muscles, weak muscles, paralysis, twisting movements, stormy movements, and tremors to mention a few of them.

Fig. 1 Cerebral palsy [5–7]

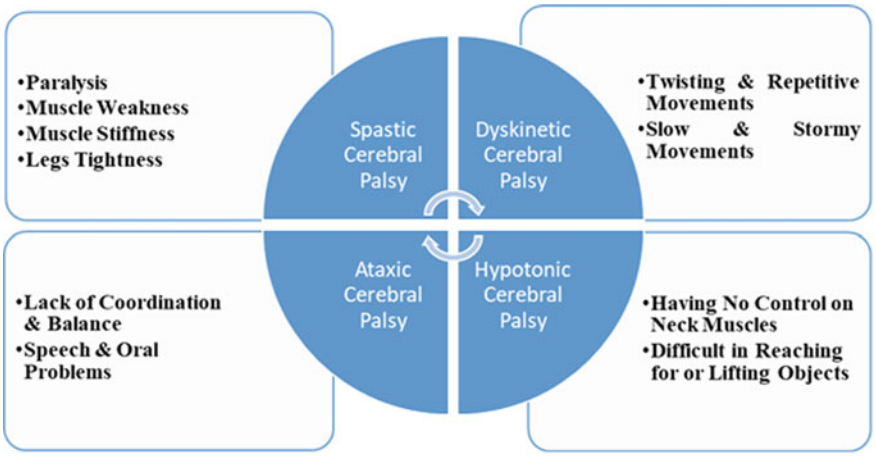


Fig. 2 Classification of cerebral palsy [8–11]

There are mainly four categories of cerebral palsy such as spastic, dyskinetic, hypotonic, and ataxic cerebral Fig. 2, due to which different symptoms are exhibited.

The organization of this paper is as mentioned: Sect. 2 discusses the spastic cerebral palsy. Section 3 discusses various techniques to examine spastic CP. Section 4 discusses the broader categories of available techniques. Section 5 concludes this study on CP.

2 Spastic Cerebral Palsy

Spastic CP is the predominant category in infants which is shown in Fig. 3. Infants with a spastic cerebral palsy struggle to eat, speak, and control their movements. Many infants with a spastic cerebral palsy walk with an abnormal gait, such as eating, speaking, and walking on their toes instead of flat feet. Symptoms of spastic cerebral palsy are as follows:

1. Stiff, tight muscles (hypertonia) on one or both sides of the body,
2. Problem of joint's extension,
3. Exaggerated movements,
4. Abnormal reflexes,
5. Limited mobility,
6. Walking on tiptoes,
7. Abnormal gait,
8. Crossed knees,
9. Contractures.

Fig. 3 Spastic cerebral palsy [12]

Spasticity is the chief symptom in most infants suffering from cerebral palsy. It is explained as a motor dysfunction due to upper motor neuron syndrome (UMNS) causing an increase in tonic stretch reflexes. Cerebral palsy infants encounter various troubles such as difficult independent ambulation due to abnormal posture gait, and they also have a joint impairment and in extreme cases, pain, and tenderness. Spasticity of lower limbs typically accompanies clonus, involuntary muscular spasm involving repeated, often rhythmic, contractions and relaxations in muscles [13]. Treatment for spastic cerebral palsy varies with an individual case. The severity of signs, area of movement problems, and secondary situations are the biggest elements in charting treatment. However, there are few fundamental methods of treatment for CP which include physical therapies, speech therapies, occupational therapies, medications, and surgery.

3 Techniques to Examine Spastic Cerebral Palsy

Pediatricians measure the angle manually or using goniometer as shown in Figs. 4 and 5 that does not give accurate angles in infants to predict spastic cerebral palsy. The evaluation of spontaneous "general movements" is a key examination of cerebral palsy in children, i.e., locomotion made in lying down position on a flat surface while not surrounded by any things that inhibit movement or tempt mind [14]. Although general movements assessment (GMA) can predict cerebral palsy with excessive accuracy, it is not broadly followed clinically because pediatrician should be trained precisely in assessing movement patterns and individuals' judgments are subjective too.

Fig. 4 Manual examinations by pediatrician [1]

Fig. 5 Angle measurement using goniometer [24]

Pervasive attempts have been made to collect quantitative data regarding the angle of the limbs which is decisive information in detecting CP. With the technical development, clinicians easily do the documentation of the electromyography (EMG) and kinematics assessment of individuals. Gait patterns have also been analyzed to achieve quantitative information. Moreover, in recent years, gesture recognition using worn sensors has become famous due to extended processing ability on mobile gadgets and the increasing accessibility of software, complicated sensors, and algorithms [15].

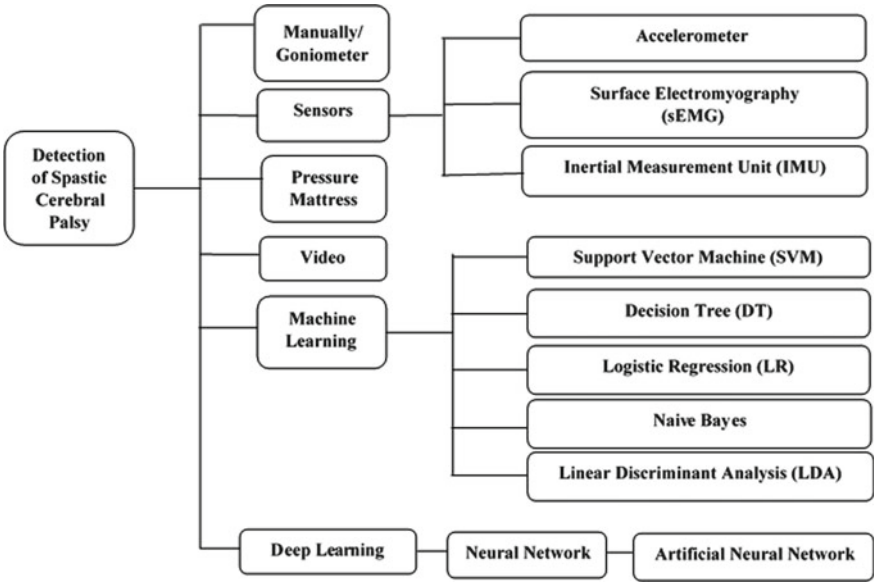


Fig. 6 Detection of spastic cerebral palsy using different techniques

In the past couple of years, many advance machine learning techniques have been developed to classify cerebral palsy infants and normal infants with high accuracy, for example, unsupervised learning methods like decision tree classifiers. Many of these strategies had been exploited in the clinical-related applications which include gait analysis. Out of many machine learning techniques, particularly support vector machine (SVM) and kernel Fisher discriminant (KFD) are known as high-performance classifiers, with the higher dimensional input dataset. In the era of technology, digital cameras and optoelectronic systems are also employed to detect abnormality of movement in infants. All such techniques to detect spastic CP in infants can be categorized in various categories as shown in Fig. 6.

During the last three months of fetus life, cerebral maturation happens rapidly. Cerebral maturation brings about continuous modification of muscle tone. For detecting cerebral palsy, the examination of postural reactions of muscle tone and reflexes is important [16]. Resting posture or attitude can be categorized into passive tone and active tone. The passive tone indicates the extensibility of a muscle when physicians apply positive actions to the infant at rest in a passive position so that the amplitude of passive kinesics of a single joint can be measured. On the contrary, the examination of an active situation in a child known as active tone, for instance, a pediatrician should place a child vertically in the righting reaction of the trunk as shown in Fig. 7.

Authors in [5] detect the lower limbs' assessment of the MTS improves the accuracy and reliability of the MTS itself. In this study, a control group of a total of thirty subjects including a CP group of twelve CP infants, a TD group of eight

Fig. 7 Righting reaction [17]

typically developed children, and an AD group of ten healthy adults were taken. Sixteen sEMG sensors have been gathered bilaterally from eight muscular tissues of the lower limb of the subjects for the duration of the ahead walk at a comfortable velocity. Surface electromyography signals of lower limb muscle tissue and acceleration records are compiled concurrently with the usage of a homemade multi-channel system consisting of sEMG sensors. Authors opine that greater than eight muscular tissues can be taken into consideration to enhance the accuracy of consequences for the prediction of cerebral palsy.

Movement estimation in infants has been done using a video/inertial measurement unit hybrid system [3]. Based on this movement estimation, a classification accuracy of 84% was achieved in the diagnosis of abnormal locomotion in infants using SVM. Twenty infants between the 2 and 4 months age were recruited for this hybrid system. The authors used five *Shimmer 3* sensors, attached on the left and right ankles, left and right wrist, and torso using soft wrist bands and ankle bands as shown in Fig. 8. Medical experts analyzed the fidgety and non-fidgety movements in infants by recording video. The results in detecting fidgety locomotion are giving good accuracy on a small dataset, but it needs to be confirmed on a large dataset too.

A combination of five IMUs along with two pressure mattresses has been used for the infant's arm motion and trunk posture assessments [2]. Firstly, this combination is applied to a human replica that has realistic anthropometric characteristics of newborns. A human replica fitted with five IMU bands, one on each upper arm, one on each forearm, and one on the torso is positioned on top of two pressure mattresses as shown in Fig. 9. After that, motor pattern parameters in a healthy child are acquired by keeping the baby on the pressure mattresses along with one torso, two forearms, and two upper arms IMU sensors to confirm the suitability of method and parameters. Using this setup, motor pattern parameters inaccuracy under ten percentage and kinematic estimation inaccuracy in a range of two centimeter are achieved. So an acceptable accuracy is achieved with a simple to use setup.

An IMU-based pose estimation method using prolonged Kalman filter and kinematic chain modeling is adapted in this approach [18]. IMU-based pose estimation has been used to analyze and reduce the sensitiveness to sensor placement and body span. This model is used for lower body (hip, knee, and ankle) pose estimation during

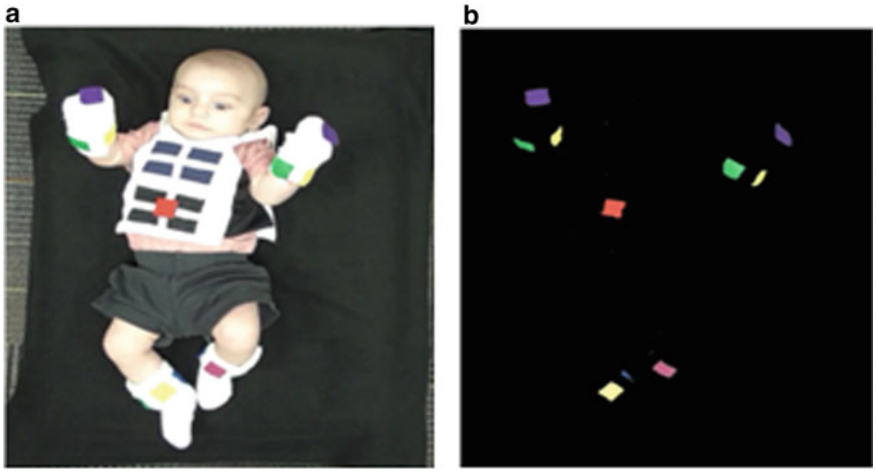


Fig. 8 **a** Infants with the shimmer sensors and color patches [3]. **b** Detection color plot [3]

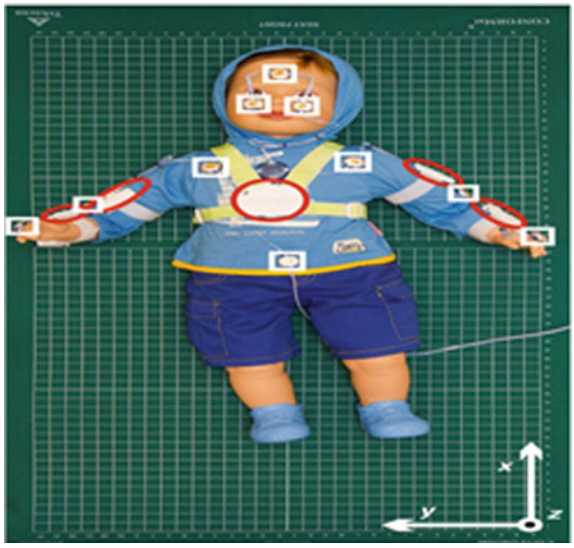


Fig. 9 IMU sensors attached on upper limbs and trunk [2]

clinical movement tests such as the single-leg squat (SLS), and the sensitiveness to parameters in addition to the knee and ankle calibration is investigated. A simple calibration protocol figuring out the IMU orientation on the body can provide better posture estimation performance. The major benefit of this approach is that it provides a tri-axial estimation of the knee, ankle, and hip joints together.

Cerebral palsy is predicted by gesture recognition in infants [15]. These gestures, known as cramped-synchronized general movements (CSGM) especially correspond with a prognosis of CP. The authors recorded information from ten infants admitted

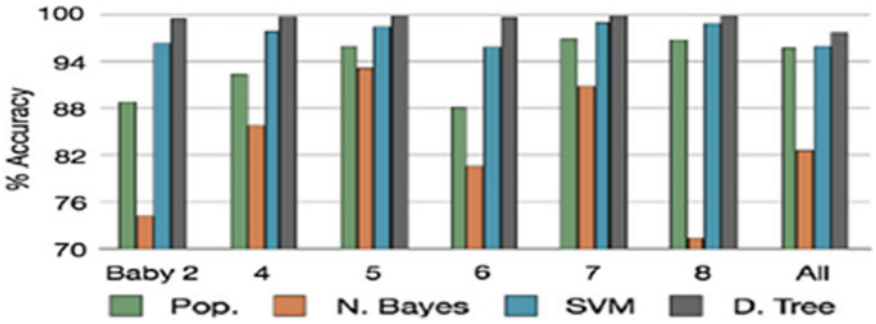


Fig. 10 Accuracy of different machine learning techniques [15]

to the new child extensive care unit at UCI Medical Center. In this method, five accelerometers are attached to an infant's upper limbs (two on the wrist) and lower limbs (two on ankle) and one on the trunk. Machine learning techniques, for instance, decision tree, SVM, Naive Bayes, and PCA, are applied on a dataset of ten infants. Results from this test exhibit that decision trees acquire the greatest accuracy of 99.46% among all the techniques tested succeeded by SVMs (90.46%). Naïve Bayes carried out worse (70.43%) as shown in Fig. 10.

Identification of gait (on foot manner of the individual) is essential for analysis in addition to for the right evaluation of the treatment consequences. Authors in [19] discover the utility of SVM for automated finding and classifying the CP based on two essential spatial gait parameters (stride duration and cadence) as input aspects. A total of 156 children dataset having 68 normal healthy infants and 88 with spastic diplegia from cerebral palsy children's dataset is used in the SVM method. The results of examination of a dataset using a tenfold cross-validation scheme revealed that an SVM classifier classified the infant groups with an accuracy of 83.33%. To get more certainty in the result, the gait parameters like an individual's lower limb length and age were added. The total dataset of 156 infants was divided into ten equal-sized segments in which six segments consist of 16 infants and the rest four segments consist of 15 infants. The accuracy of SVM results is higher than linear discriminant analysis (LDA) and a multi-layer perceptron (MLP) based classifier which is 3.21% and 1.93%, respectively. SVM kernel functions such as linear, polynomial, radial basis, and ANOVA spline were studied in this approach. Among these four functions, the polynomial and radial basis kernel performed better than the other two functions.

Many known classification methods have been carried out in numerous applications such as decision trees, Fisher's linear discriminant analysis (FLD), and artificial neural networks to name a few. Also, many methods have been used to examine the performance of pattern classification schemes. Generally, holdout, leave-one-out, and cross-validation methods are used for evaluation. Authors in [20] explore various classification paradigms in the cerebral palsy gait analysis. In this technique also, a dataset of 68 healthy infants and 88 infants with spastic diplegia is used with parameters such as stride length and cadence. The tenfold validation method along

with kernel Fisher discriminant analysis (KFD) for the classification of CP gait has been used in this approach too. Many different cross-validation consequences show that KFD offers better category accuracies than SVM and is advanced to a few other class methods consisting of the decision tree, multiple layer perceptron, and *k*-nearest neighbor.

Wearable sensors such as accelerometer, sEMG, and Opal to mention a few are used to evaluate an infant's limb moves the whole day. Authors in [5] used acceleration and angular momentum to evaluate infant-produced movements. Two accelerometers are attached to the infant's lower limb across 8.5 h, during which the child is taken care of by the nurse. Accelerometers are used in this study to track acceleration only and recording the leg movement. But in this approach, around fifteen percentage of child leg moves can be credited due to the background motion of the nurse. Table 1 provides a comparison among the various cerebral palsy detection techniques.

Table 1 Comparative analysis of various CP detection techniques

S. No.	Technique	Measurement	Pros	Cons
1	Manually/goniometer	Joint angle	Identify active and passive tone	No accuracy
2	Video	Limb movements	Detecting fidgety movement	Dependency for capturing video
3	Accelerometer	Limb movements	Recognize the gesture	Measures only horizontal and vertical orientation
4	sEMG	Limb movements	Easy to understand muscles abnormality Record single muscle activity	Only feasible for single muscle in an individual
5	Pressure mattress	Infants' activity level	Simplicity to use	No joint angle measurement
6	IMU (accelerometer + gyroscope)	Joint angle of both limbs (orientation and rotation both)	Accurate angle Low cost Non-invasive	Magnetic field cannot be measured
7	IMU (accelerometer + gyroscope + magnetometer)	Joint angle of both limbs (orientation, rotation, and magnetic field also)	Accurate angle Low cost Non-invasive Measures three-dimensional data on each of three-axis	–
8	Machine learning	Stride length, cadence, leg length, and age	More accurate results	Larger dataset is required

Another big research in cerebral palsy is done using multivariate analysis and machine learning techniques. Multivariate analysis is useful for the prediction of cerebral palsy and machine learning approaches made it viable to automatically identify movement impairments in high-risk infants. It is anticipated that multivariate analysis and machine learning will play a massive role in improving the investigation and remedy of cerebral palsy to diminish fatality and morbidity rates and enhancing patient care for infants with CP [6]. The multivariate analysis includes statistical techniques, for instance, principal component analysis (PCA), canonical variate analysis, independent components analysis, and multivariate regression.

Researchers have used different sensors such as accelerometer, sEMG, pressure mattress, and IMU. These sensors are typically attached to upper limbs and trunk to measure the angle of joints. In some literature work, machine learning techniques, such as SVM, decision trees, and principal component analysis based on four parameters: (i) stride length, (ii) cadence, (iii) leg length, and (iv) age, are also used to predict or classify CP in infants. All these techniques individually have their specific limitations such as the machine learning technique needed more data for accuracy. A combination of sensor-based and machine learning techniques overcomes these limitations and provides a better accuracy to detect/classify CP in infants.

4 Discussion

In the area of neurodevelopment related to the detection of cerebral palsy, big tries had been made to gather quantitative information about the angle of the limbs which is decisive information. Pediatricians measure the angle manually or using goniometer as shown in Figs. 4 and 5. Sometimes pediatricians could not measure the angle accurately in an infant which might lead to the wrong detection of CP. Nowadays, cerebral palsy is detected using different techniques such as sensor-based, machine learning, and video-based approaches employing specific parameters as mentioned in Table 2. These techniques provide much better accuracy in comparison with the conventional

Table 2 Parameters used in various techniques to detect cerebral palsy in infants

Techniques	Parameters to detect/classify CP
Sensor-based 1. Accelerometer 2. sEMG 3. IMU	1. Angle between joints of upper and lower limbs
	2. Movements of the infant
Machine learning	1. Stride length
	2. Cadence
	3. Leg length
	4. Age
Video-based	1. Movement of the infant's upper and lower limbs

measurement method by goniometer. Using the sensor-based techniques, pediatricians can measure the angle between joints of the upper limbs, lower limbs, and trunk accurately. Sensors such as Opal, accelerometer, sEMG, and IMU to name a few are used to measure the angle between joints that provide the 3D image of the limbs. By this image, a pediatrician can predict the CP early in infants.

Existing motor evaluation techniques like digital cameras and optoelectronic systems are afflicted by material occlusion and require complicated setups [3]. Electromyography (EMG) can be used only for recording the electrical activity of muscle tissue or visual representation. Some sensors like accelerometer, sEMG, and IMU are very useful for neurodevelopment assessment [21]. Wearable sensors like accelerometer have turn out to be increasingly famous as a measurement of movement aspects and physical actions. Accelerometers have been used to observe neonate's continuous upper and lower extremities moves however do not offer postural facts. To overcome these drawbacks, multi-sensor measurement systems provide a better alternative to detect the motor fidgety movement. IMU sensor as shown in Fig. 11 is the combination of three sensors, namely accelerometer, gyroscope, and magnetometer which is shown in Fig. 12. From the survey carried out, we found that the IMU sensor gives a more accurate angle between joints as compared to other sensors because it measures orientation, rotation, as well as the magnetic field.

Fig. 11 IMU sensor [22]

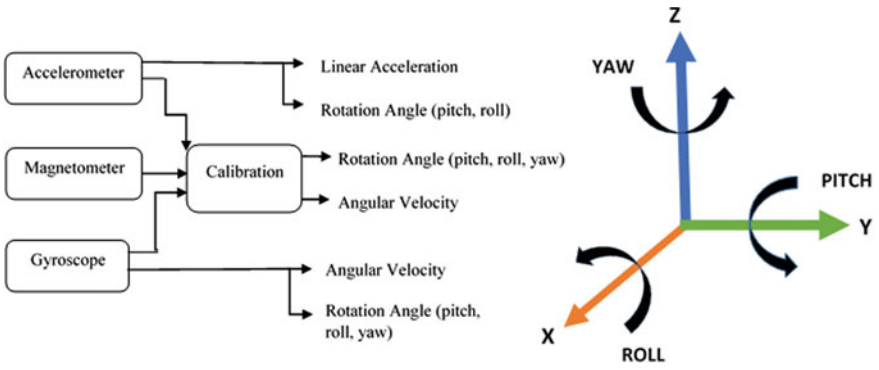


Fig. 12 IMU comprising of three types of sensors and working principal [21]

4.1 Accelerometer

An accelerometer is an electromechanical tool used to measure acceleration forces [23]. An accelerometer measures the movements of the limb and quantifies them such that each break or modification is counted as a new movement.

4.2 Surface Electromyography (sEMG)

Electromyography (EMG) can be used only for recording the electrical activity of muscle tissue or visual representation. sEMG assesses muscle function by recording muscle activity from the surface above the muscle on the skin [24]. As such, sEMG is the analytic device that is used in a broad spectrum that assesses muscular coordination in cerebral palsy.

4.3 Inertial Measurement Unit (IMU)

Wireless IMUs shown in Fig. 11 are low cost, non-invasive, and easy to wear. This sensor is the combination of a tri-axial gyroscope, a tri-axial magnetometer, and a tri-axial accelerometer [25]. IMU measures the angle of joints and movements when an infant moves his/her limbs. It can measure movement throughout 360° using three sensors each on three axes and hence used to analyze infants' spontaneous upper and lower extremity movements. There are three types of rotation angles: pitch [rotations of the *y*-axis (±90°)], roll [rotation of the x-axis (±180°)], and yaw (rotation of the *z*-axis (±180°)] as shown in Fig. 12.

The benefit of the usage of IMU is that it will no longer be interfered with by using the external magnetic area across the sensor while used near the ferromagnetic material. Using a combination of accelerometer and gyroscope handiest, it may not be enough to increase the accuracy of the measurement because of sensors' noise and the gyros drift problems. So magnetometer measures the magnetic field when a force is applied [21].

IMU sensor is used in many fields such as augmented reality, health care, drone, navigation system, and robotics. For example, an IMU having only a gyroscope and an accelerometer was used in a device to study the movements of the post-traumatic patients [21]. This IMU records a three-axis accelerometer and/or three-axis gyroscope data at many samples per second, permitting to file movement records unobtrusively across many continuous hours without any intervention.

4.4 Machine Learning

Many machine learning algorithms such as SVM, decision tree, and PCA have been applied for detecting CP based on infants' limb movements [15]. Machine learning techniques have been applied only for the classification of healthy infants and CP infants. But none of these techniques has been used for predicting cerebral palsy in infants at early stages so far. Four parameters (1) stride length, (2) rhythm, (3) lower limb length, and (4) age are taken in machine learning techniques to classify the CP infants and healthy infants.

4.5 Hybrid System (Video and IMU)

Some authors also have used a video and the IMU hybrid system for assessment of the infant's movement. In this hybrid system, IMU sensors are attached to infants' limbs. After that, Infants' movements are recorded in a video. From the video recording, pediatrician can classify the healthy infants and CP infants easily based on the abnormal movements. Cramped-synchronized general movements (CSGMs-abnormalities) identified by a trained specialist from the video have been correlated with CP.

5 Conclusion

The diagnosis of cerebral palsy using a machine learning technique is more preferable. Parameters such as age, stride length, leg length, and cadence are used in machine learning techniques to detect cerebral palsy in an infant. The IMU-based technique is very useful to detect cerebral palsy in infants at early stages. No research has been done on the IMU-based technique applying on all the limbs. IMU sensor can be used on the trunk, upper limbs, as well as lower limbs for measuring angles between arms, ankle, and other joints. Thereafter, a machine learning model can be developed based on the measurements taken using IMU to detect whether the infants are suffering from cerebral palsy or not. Based on the outcomes, the model can be validated against the diagnosis which is done by the pediatricians. This study concludes that the IMU sensor gives better accuracy for angle measurement.

References

1. https://www.healthynewbornnetwork.org/hnn-content/uploads/India-1.pdf.
2. Rihar, A., Mihelj, M., Pašič, J., Kolar, J., & Munih, M. (2014). Infant trunk posture and arm movement assessment using pressure mattress, inertial and magnetic measurement units (IMUs). *Journal of NeuroEngineering and Rehabilitation, 11*(1), 133.
3. Machireddy, A., Santen, J. V., Wilson, J. L., Myers, J., Hadders-Algra, M., & Song, X. (2017). A video/IMU hybrid system for movement estimation in infants. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).*
4. https://en.wikipedia.org/wiki/Cerebral_palsy.
5. https://pediatrics.aappublications.org/content/130/1/e152/tab-figures-data.
6. Zhang, J. (2017). Multivariate analysis and machine learning in cerebral palsy research. *Frontiers in Neurology, 8,* 715.
7. https://kimpediatrics.com/march-national-cerebral-palsy-awareness-month/.
8. https://www.cerebralpalsyguide.com/cerebral-palsy/types/spastic/.
9. https://cerebralpalsygroup.com/cerebral-palsy/ataxic/.
10. https://kidshealth.org/en/parents/dyskinetic-cp.html.
11. https://www.cerebralpalsyguidance.com/cerebral-palsy/types/hypertonic/.
12. https://in.pinterest.com/pin/301741243755840638/?lp=true.
13. Choi, S., Shin, Y. B., Kim, S. Y., & Kim, J. (2018). A novel sensor-based assessment of lower limb spasticity in children with cerebral palsy. *Journal of Neuroengineering and Rehabilitation., 15*(1), 45.
14. https://neurologicexam.med.utah.edu/pediatric/html/newborn_n.html.
15. Singh, M., & Patterson, D. J. (2010). Involuntary gesture recognition for predicting cerebral palsy in high-risk infants. In *International Symposium on Wearable Computers (ISWC) 2010.*
16. Amiel-Tison, C. (1968). Neurological evaluation of the maturity of newborn infants. *Archives of Disease in Childhood, 43*(227), 89–93.
17. https://www.indianpediatrics.net/june2006/june-527-534.htm.
18. Kianifar, R., Joukov, V., Lee, A., Raina, S., & Kulić, D. (2019). Inertial measurement unit-based pose estimation: Analyzing and reducing sensitivity to sensor placement and body measures. *Journal of Rehabilitation and Assistive Technologies Engineering, 6,* 205566831881345.
19. Kamruzzaman, J., & Begg, R. (2006). Support vector machines and other pattern recognition approaches to the diagnosis of Cerebral Palsy gait. *IEEE Transactions on Biomedical Engineering, 53*(12), 2479–2490.
20. Zhang, B.-L., & Zhang, Y. (2008). Classification of cerebral palsy gait by Kernel Fisher Discriminant Analysis. *International Journal of Hybrid Intelligent Systems, 5*(4), 209–218.
21. Ahmad, N., Ghazilla, R. A. R., Khairi, N. M., & Kasi, V. (2013). Reviews on various inertial measurement unit (IMU) sensor applications. *International Journal of Signal Processing Systems, 1,* 256–262.
22. https://www.ceva-dsp.com/ourblog/what-is-an-imu-sensor/.
23. https://www.livescience.com/40102-accelerometers.html.
24. https://en.wikipedia.org/wiki/Electromyography.
25. https://simplifaster.com/articles/buyers-guide-imu-sensor-devices/.

Data Security Using Directory Server in Identity and Access Management System

Manav A. Thakur, Thaksen J. Parvat, and Vishal S. Walunj

Abstract Data is crucial in today's word. Each and every company function depends on data. Depending upon the data storage capacity of companies, it is segregated into various categories like multinational, small-scale, and middle-scale companies. As an example, Facebook, Google, Amazon, etc., they have billions of users and their data, which are continuously uploading/downloading from their Web sites. Data is the core part of every company working. To store the data, companies are using different database and directory servers. The databases or directory servers store the data in their won formats and to extract the data from this, and they are using special functions and protocols. Data is mostly used to manage the identity and access of the users into the organization and outside the organization. IAM is a suit which nowadays very popular in the market for securing the user data. IAM system allows the user to access the data depending upon user credentials, which means when the data is accessible by the user. The user has to enter the credentials, and then depending upon credentials, the access is granted to the user. The IAM system proposed in this paper depends upon three things (WHO, WHAT, and HOW) Who's identity can be created and who can access the data, also the level of data access will be decided and How the user can access the data. In this using the single directory server, we will try to implement a single server system in that both identity and access manager can use the same data storage but having a different suffix. IAM suite will also make sure that users can access the data for which they have the access and need for their business operations.

Keywords Directory server · LDAP · Identity management system · Identity manager adapters · Access management system · Single sign-on (SSO)

M. A. Thakur (✉) · T. J. Parvat
D. Y. Patil College of Engineering and Innovation Varale, Pune, India
e-mail: manavthakur84@gmail.com

T. J. Parvat
e-mail: thaxcen@gmail.com

V. S. Walunj
CSMU, Panvel, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_8

1 Introduction

From the past many years, scientist and network administrators have worked hard for developing such secure terminology, method to secure user data. But still users are not feeling secure while using any payment gateways, e-commerce Web site, and business applications. So to remove insecurity from user mind, nowadays many companies are moving toward the identity and access management (IAM) system, which helps them to secure the organization environment.

In many companies, top management and system administrator nowadays are putting their complete efforts toward user data protection, user data compliance for auditing purpose, and data security-related cases. This more focus on user data security or data auditing and compliance will help them to secure managed system into the organization, which automatically increase shareholder value into the organization. This data auditing and compliance will help us to avoid negligence into user data security, which will help to take toward the best result of security. Looking from a data security perspective, it is always good practice to show more care toward achieving well-planned security system, exactions of system policies, proper control over the process, and achieve the best result of system security. Using the modern identity and access management (IAM) system designing and implementing role base and access base privileges of individual users on the network is possible. In the IAM system, users are granted (or denied) these role base and access privileges. These users can be customers of the company or employees of the company. The core objective of IAM systems is that the user is holding only one digital identity. Once the user digital identity has been created into the directory server under the identity management suffix, that digital identity is maintained, monitored, and can be modified when required by the users through the user "lifecycle management." Such secure system will help to achieve the main goal is to manage the user identities, user authentication, authorization, increase user security, manage organization productivity, secure user access privileges, manage user roles, and user organizational groups.

But the typical IAM system which is available in the market is using the different data storage to store the data; this will lead them to use different directory server for identity management and access management. Because of different data storages, there is network complicity increase, need an extra server, cost increase, error identification difficult, need extra help desk resource (Figs. 1 and 2).

So to avoid such problems, we can use the single directory server for identity management and access management. In our proposed system, we are using single data storage but with different suffix as an example, cn = identity management and cn = access management (suffix). It will help us to manage the system easily. In identity management, suffix identity-related data will be stored, and in access management suffix, user authentication and authorization data will be stored. We can also do the replication of data if required. This concept will help to reduce the system complexity, decreasing cost, administrative work, and downtime of servers. Many companies are providing such IAM systems.

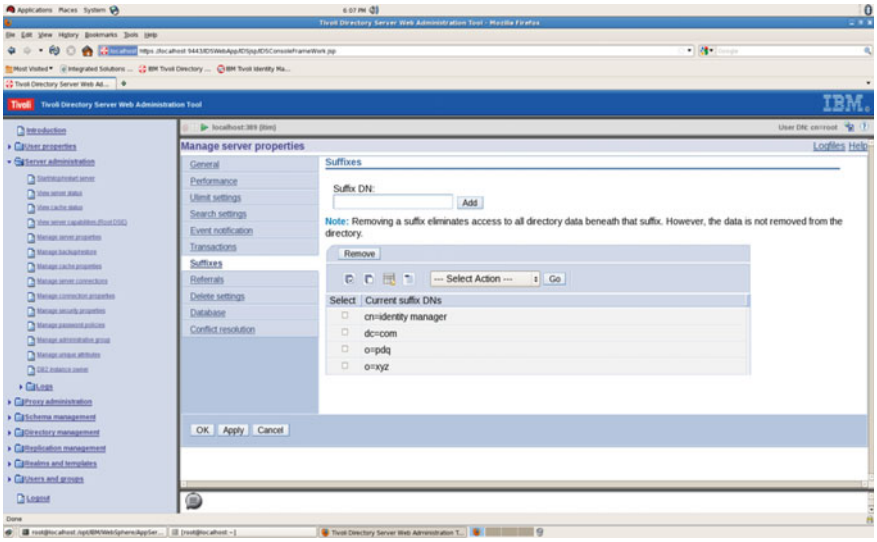


Fig. 1 Separate suffix for identity management

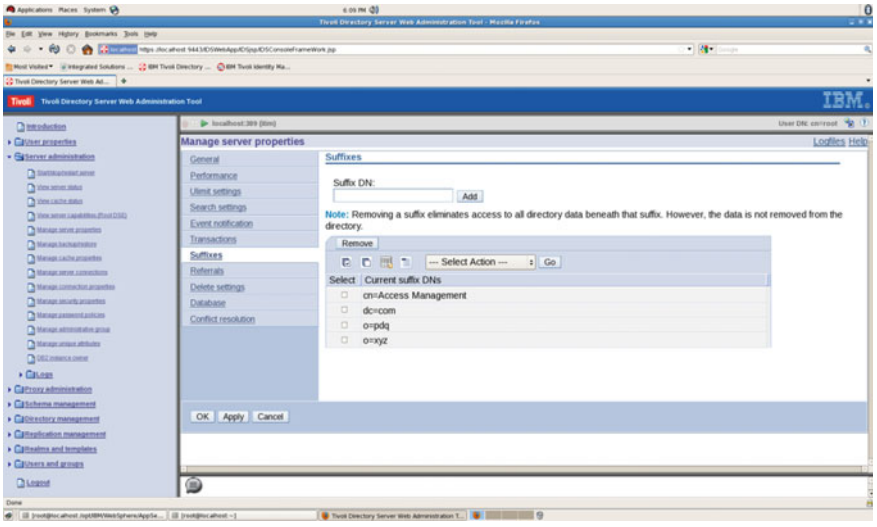


Fig. 2 Separate suffix for access management

- Microsoft Azure,
- IBM IAM SYSTEM,
- Okta's identity and access management,
- Oracle Identity Cloud Service,
- Centrify Identity Service,

- RSA SecurID Suite (Dell company),
- Chicago-based Keeper Security,
- SailPoint IdentityIQ (Austin, Tex.-based),
- OneLogin (San Francisco-based),
- Ping (Denver-based).

2 Directory Servers

In today's world, many storage devices are available to store the data. Nowadays, the data is available on the tip of our fingers by one click, as there are many software and hardware products which help them to store, upload, and download data easily. One such accessible storage that is available in the market due to its fast data retrieval capacity is directory server. A directory is mostly considered as a database server. Still, it is a specialized database having unique characteristics that make it apart from general-purpose relational databases [1]. Directory server is the centralized data storage used in today's network security infrastructure. Data mapping between network resources and network addresses using directory server uses two ways:

- Lightweight Directory Access Protocol (LDAP)
- Directory Services Markup Language (DSML).

In the preceding Fig. 3:

- Here, dc is noted as domain component
- ou is organizational unit
- cn is surname.

In directory, server data is formatted in the form of Distinguish Name (DN) and Relative Distinguished Name (RDN). It is called the tree structure of the directory server. This tree structure is starting into the directory server from suffix. In the IAM system, directory server plays a significant role; it acts as a central repository for data storing and managing user information. Directory server almost stores all types of information, like users' identity profiles, access privileges, and information about an

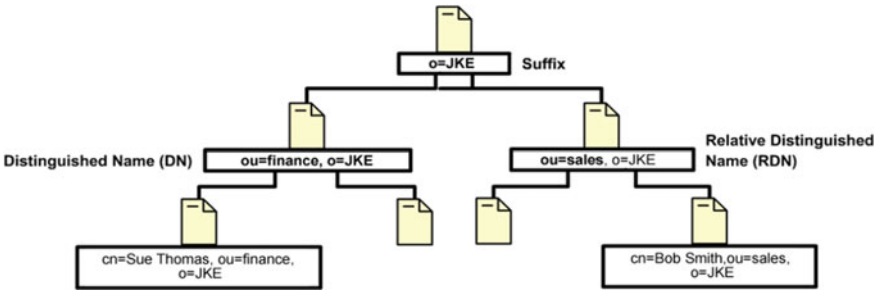


Fig. 3 Directory information tree structure

application, network resources, printers, network devices, and manufactured parts. IAM systems perform authentication and authorization of users using the directory server [2]. It is nothing but the backend server from where the user is getting validated. Depending upon the authentication result of the user, the user is authorized and will get the access to backend application. The advantage of using the directory server is that it can be easily integrated with any available data storage into the organization. Once integration is done, then it exchanges the information of already available user, customer, and business partner [3, 4].

2.1 LDAP: Protocol

The Lightweight Directory Access Protocol (LDAP) is a message protocol used to connect between directory servers and directory clients. The LDAP uses different methods. For example, when the client wants to access the data from the server at the beginning of a connection, a client bind request raised to the LDAP server and then a search request is used to search data from the directory server. All this client–server communication of LDAP executes using TCP/IP protocol [5] (Fig. 4).

The IAM system all uses identities, user access, user accounts, their groups and roles, organization chart, organization policies, backend resources, and decision and approval workflow is going to store into directory server using the LDAPv3 directory server version. In the IAM system, using the LDAP, many authentication mechanisms can be executed like simple authentication, form authentication, and token authentication mechanism [2] (Fig. 5).

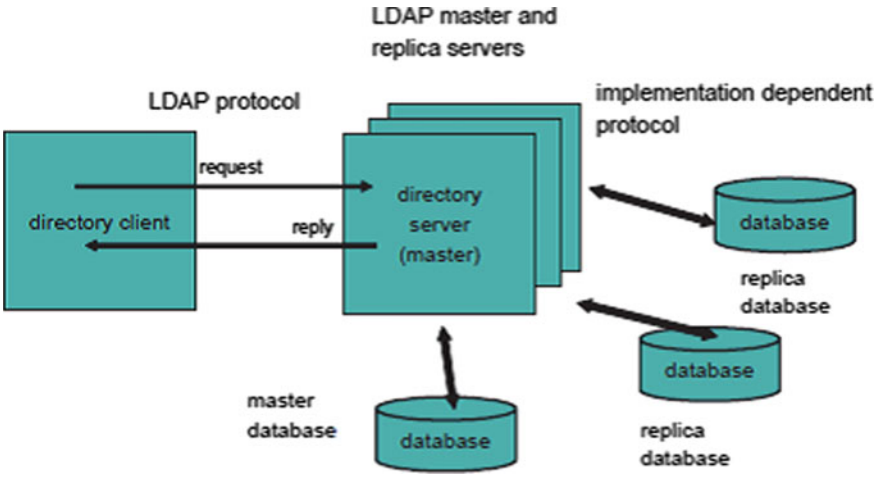


Fig. 4 LDAP client–server model

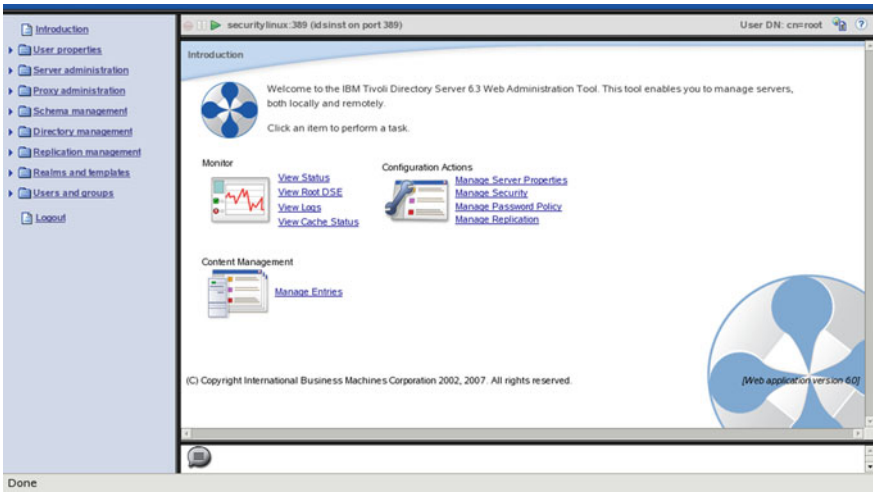


Fig. 5 Directory server administration tool

When the installation and configuration of your directory server are done with the IAM suit, many components are getting installed.

- An LDAP server for processing requests from the client. (Client–server can be on the same machine, but in high availability, we can install it on separate server or machine.)
- Administration server and administrator console to execute admin queries of Sun Java Company.
- Server-side command-line tools which are also used to start and stop the directory server, import the data into directory server, export the data when required, to active the user, and to inactive the user.
- GUI which work as front ends and responsible for communication of LDAP and DSMLv2 language.
- Access control list
- Directory server replication functionality.
- Directory schema files consist of directory tree.

3 IAM Systems

Controlling user identity and access to networks, data, and applications is one of the most crucial jobs for security team plays. The best solution to this critical challenge is to implement identity and access management (IAM) products. IAM suite help IT departments to ensure that the identity and access to the cloud, on-premises, and hybrid environments are at the right level. IAM system in the IT environment is about giving access privileges, used to define and manage the roles of user and

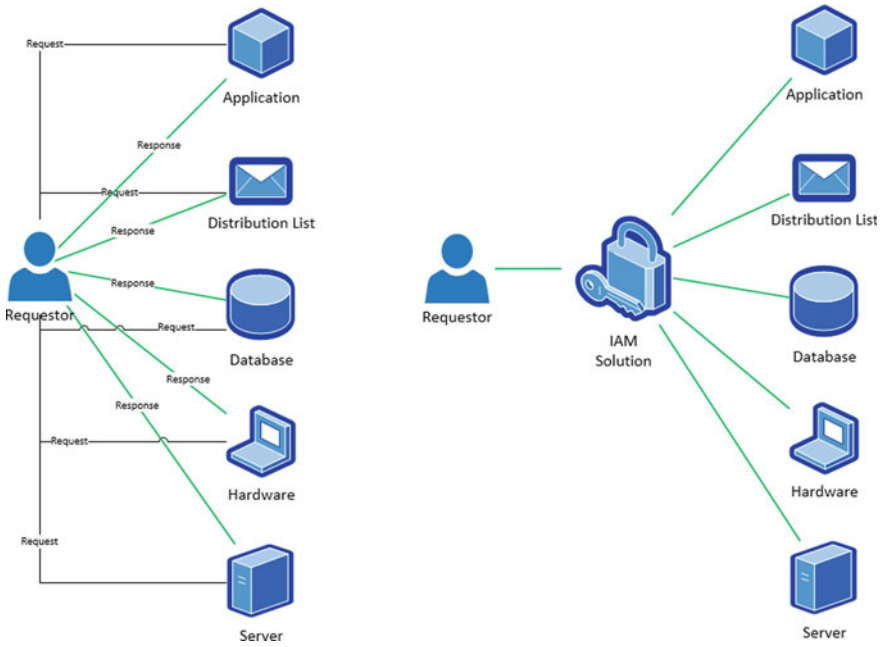


Fig. 6 IAM system basic architecture

control the access of users and the execute such condition where users are getting (or denied) those access [6]. Those users might be an employee of organization (employee identity management) or customers (customer identity management). The main objective behind IAM systems is one digital identity per user or business partner. Once the digital user identity has been created into the directory server, it has to be kept for a long period of time, when need can be continuously modified and monitored throughout complete each user's lifecycle. [7].

Today's IAM system also provides facility to on-boarding and off-boarding of users, which results to secure access to your systems and applications automatically (Fig. 6).

3.1 Identity Management System Implementation

Let us start describing the functionality of identity management from the overall field of the IAM system. Identity management basically deals with identities of user and their business partners' identities. All the data related to the identity management system is stored in the directory server. When the data is stored into the directory server, it is stored in the form of RDN and DN into the particular suffix. In that suffix,

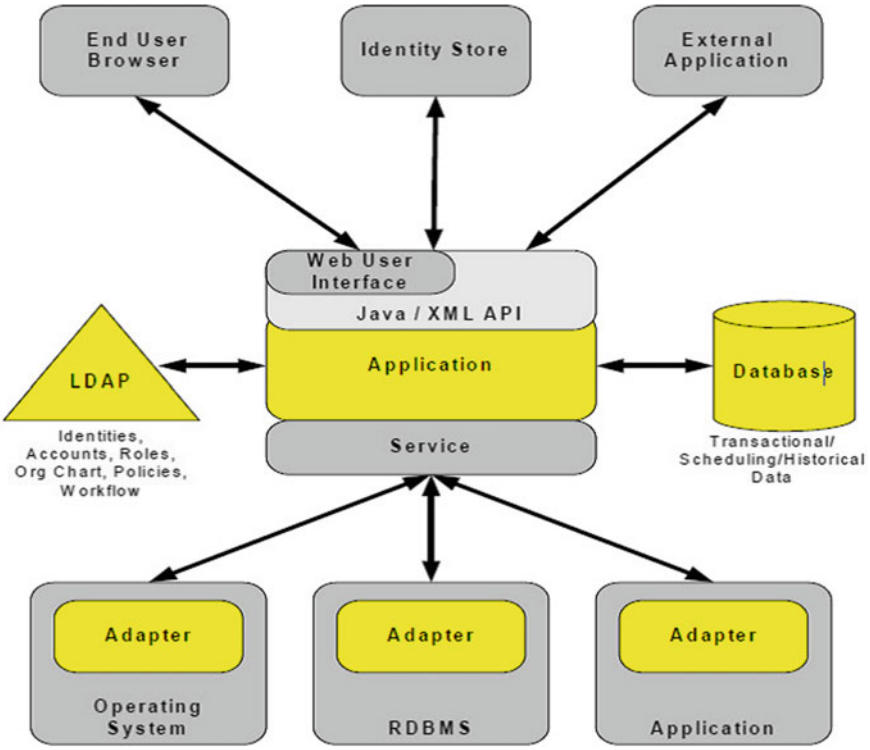


Fig. 7 Identity manager logical architecture

all the user data, policies, workflow, roles, groups access controls list and backend resources data are stored [6, 8] (Fig. 7).

The logical architecture is having basically two layers

- Presentation: The presentation layer logically represents Web User Interface functionality.
- Provisioning: This provides the provisioning facility for the users.

Along with it also consist of some other products like

- Identity Store: The identity store consists of database server and directory server (LDAP).
- Database Server: It is a relational database that stores all user transactional, reporting data for audit, and schedule information. The information available in database is stored as temporary data used to check the current state of transactions, but it also stores historical information which is used for the auditing and logging purpose of the organization.
- Directory Server (LDAP): It uses an LDAPv3 directory server to store the information. It works as the primary repository for storing the user identity and access

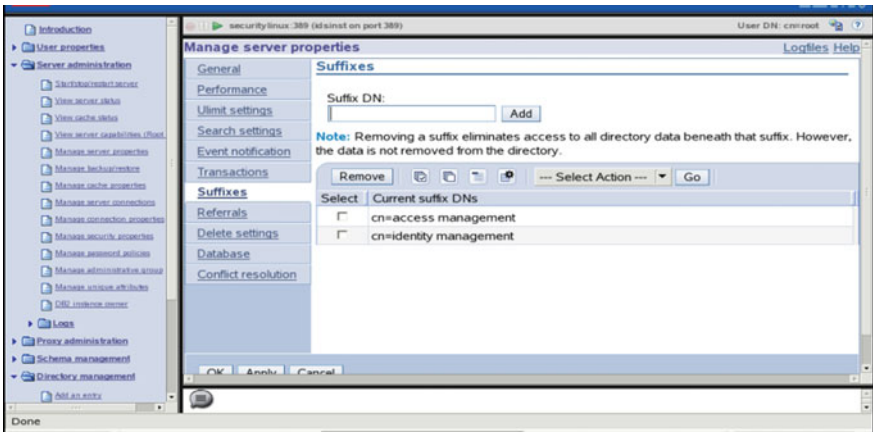


Fig. 8 Suffix details of IAM system

information. Here, in this structure, both identity and access management systems are using the same directory server with different suffix [3, 9] (Fig. 8).

- Directory Integrator: It is used to link a variety of data resources to the IAM system. Directory integrator provides the connection with data in files, directories, databases, and message queues, to manage the user identities and access. It helps to synchronize the data between resources and the IAM system. For communication with data resources, it is using the various adapters [1, 8].
- Web Sphere Application Server: It is the core component of the IAM system. It will provide the application layer server, which helps to runs the IAM application and executes all the identity management operations and help to provide the graphical user interface to identity management system.
- Identity Manager Adapters: To manage the user access and account on the resources IAM system using the special program called adapters. Depending upon the resource, there are different types of adapters used.

"Identity management" this terminology is mostly integrated with term access management. The integration of these two products into a single product is called identity and access management systems (IAMs). This provides identity and access management to the user (Fig. 9).

From many past years, there are various revolutions that happened in web technologies, the way to deliver the information, and providing services. Access management will provide functionality in the field of sales, customer service, online purchasing, cloud and various mobile applications. Due to this, the customers who are doing online purchasing and business partners who are doing sales continuously need real-time data access to the information [10].

This need or access to real-time data makes them to develop such a system which can provide them whatever the market requirement, good efficient system, cost-effective system, and real-time system which can give them advantage over the

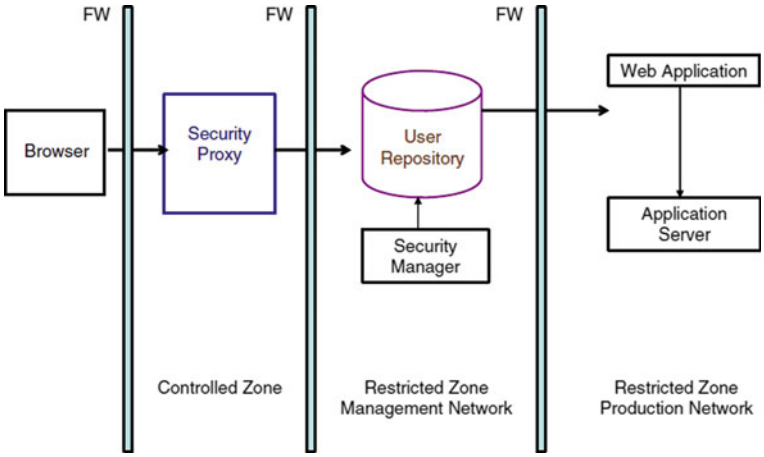


Fig. 9 Logical security infrastructure for access management

market competitors [11]. Due to this, compliance is the challenge, auditing is issue, and tracing is a problem for the companies. To make the automation of core business processes depending upon the role of users or we can say depending upon the credentials of the user, business partners, suppliers will gain the access to the company information. When any user wants to access any secure resource on the network, basically, he wants to have access to that system to perform the particular operation.

Then in such a case, the authentication process is to be executed to access information from various devices like computers, mobile applications, Tabs, or by using the location of the user using the Internet. So the first step is that the user must have his own identity, and then he should be able to login securely using the credentials from web browser or a wireless device but without taking client software into consideration. In addition, with this, there are different web services available in a large organization, among them some of web services are locally get deployed, and among them some are available on cloud technology, but when the user wants to access any service among them, user needs access privileges, for each server or application they want to access [3]. The decision to open access for IT services or backend resources is not the work of the access management system. This decision is depending upon the identity of that user in the IAM system. The user identity is available in the directory server under the access management suffix. Because the policy execution and rule-following are done by access management, this is considered during the service planning and system design. Access management is not a decision-maker, but it is enforcing the policies to restrict unauthorized users from providing access [6]. When the user identity is created under the directory server suffix or directory server tree, all his authentication and authorization are done from there when the user wants to access any protected resource. Depending upon the company policy, protected object policies (POPs), and access control lists (ACLs),

the company authorization rules are decided. In such case, directory information tree is the single point of communication from where authentication and authorization execute to access the web resources using access management [8].

3.1.1 Single Sign-on (SSO)

Single sign-on (SSO) functionality is used in identity and access management (IAM) that provides users secure authentication with available multiple backend resources without using multiple credentials. In this, the user has to login once and then available backend resources will get accessed. There are many companies like Google, LinkedIn, Twitter, and Facebook that provide SSO capability for an end-user to log into a third-party application using their social media credentials. The working of SSO solutions is depending upon the execution organization's identity store known as Directory Server [1].

4 Conclusions

This paper wants to present the use of directory server in the IAM system. Normally, companies are using different directory server for identity and access management system. That is also best practice, but in this, we use the same directory server and in the same directory server, we can put the identity and access data into two different suffixes, as shown in the diagram. This will help us to improve our performance. The directory server IAM data available in the suffix can be replicated. This replication will help us to retrieve data if original data is lost as replication is available in such a case. The use of a single directory server will also help to reduce the cost, and when we want to do the integration between identity and access management, this will help us, as data is available at a single point of server. This terminology is useful for small-scale company users. The IAM system suggested will help to reduce admin work by using the automation capability. Using the same directory server, single sign-on capability is easily integrated with access management. We can implement two-way authentication methods also.

References

1. Indu, I., Rubesh Anand, P. M., & Bhaskar, V. (2018). Identity and access management in cloud environment: Mechanisms and challenges. *Engineering Science and Technology, an International Journal, 21*(4), 574–588.
2. Gomez, S. R., Jero, S., Skowyra, R., Martin, J., Sullivan, P., Bigelow, D., & El, Z. (2019). *Controller-Oblivious Dynamic Access Control in Software-Defined Networks*. IEEE Xplore: August 22, 2019.
3. Manolache, F. B., Evans, J., Rusu, O. (2018). *Mycros—an automated enterprise IT management system based on LDAP*. IEEE Xplore: November 1, 2018.
4. Thakur, M. A., & Gaikwad, R. (2015). User identity & lifecycle management using LDAP directory server on distributed network. In *IEEE Conference 2015.*
5. Manolache, F. B., & Rusu, O. (2018). *An automated enterprise IT management system based on LDAP*. IEEE.
6. Tariq Banday, M. (2018). *Directory services for identity and access management in cloud computing*. IEEE Xplore, June 21, 2018.
7. Xu, J., & Xue, K. (2020). *An identity management and authentication scheme based on redactable blockchain for mobile networks.* IEEE, April 8, 2020.
8. Thakur, M. A., & Gaikwad, R. (2015). User Identity and Access Management trends in IT infrastructure—an overview. In *IEEE Conference 2015.*
9. Chang, C. C., & Lee, C. Y. (2011). A secure single sign-on mechanism for distributed computer networks. *IEEE Transactions on Industrial Electronics, 59*(1), 629–637.
10. Indu, I., & Rubesh Anand, P. M. (2016). *Identity and access management for cloud web services.* IEEE Xplore: June 13, 2016.
11. Torres, J., Nogueira, M., & Pujolle, G. (2012). A survey on identity management for the future network. *IEEE Communications Surveys & Tutorials., 15*(2), 787–802.

Deep Learning Techniques in Image Description

Prateek Gupta and Archana Singh

Abstract The current era of high end GPU with the exploring set of data requires in-depth analysis using deep learning techniques with the combination of computer vision and natural language processing has attracted huge attention of IT industry in the recent past. The domain of Image captioning is a great revolution with respect to deep neural networks. In this paper, the hybrid model is proposed to retrieve Image description from the dataset. The motivation behind this work to explore the deep learning techniques in image dataset which is quite impressive, although to implement the processing of different models was quite challenging to get the better accuracy of results. The paper uses the hybrid algorithm CNN-CNN and CNN-RNN based model to describe the images and visual experience of that images in the form of meaningful sentence.

Keywords Deep learning · CNN-RNN · Image · Natural language processing · Image caption

1 Introduction

In the recent years, image processing has made huge improvement, like image classification and object detection both of them have improved a lot on the basis of accuracy. The popularity of deep learning models in identifying, analysing an image with the better accuracy motivated us to propose this work using a hybrid model of neural networks techniques [1]. The work focuses on to create a description for any input image and to understand the visuals in the image thus moving in the direction of Image captioning bot. This model itself has large potential effects on our daily life such as headlines of a news coverage or can be used for blind people after converting

P. Gupta (✉) · A. Singh
Adobe Systems Noida, 2 ASET, Amity University, Noida, Uttar Pradesh, India
e-mail: pratgup@adobe.com

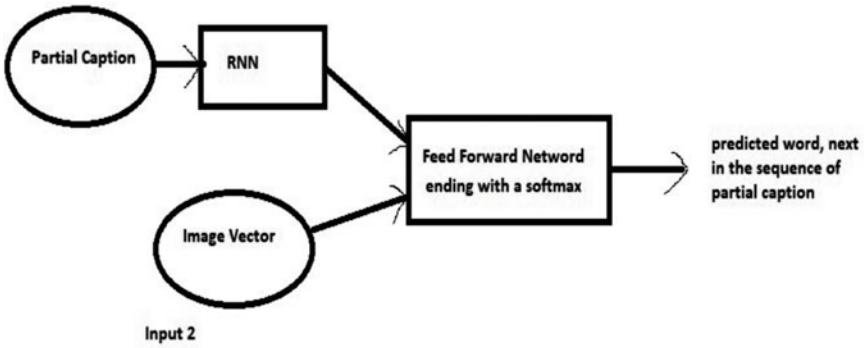
A. Singh
e-mail: asingh27@amity.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_9

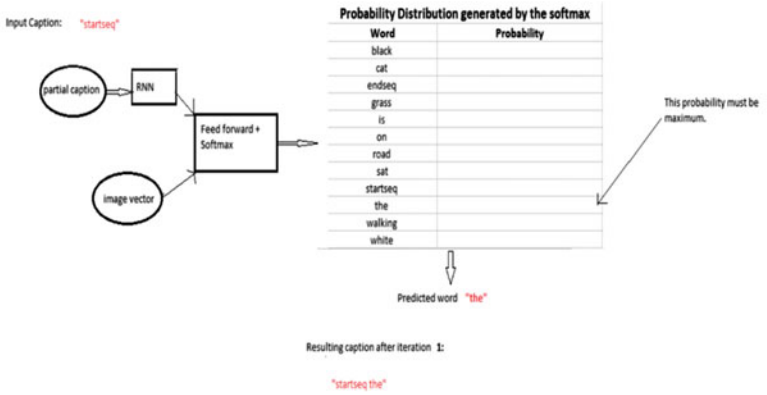
Fig. 1 An example of image captioning

to audio to let them have the aesthetic and hedonic experience in the world [2]. This topic has lots of theoretical and practical implementation in the stack which needs to be done. With input as an image, the algorithm produces a caption of the image using the lexicons of natural language. In Fig. 1, it presents the visuals in the image as it consists of a man standing on the rocky near the ocean, the proposed algorithm captures the nouns and action happening in the image. This paper focuses on the model to describe Image description, it requires the combination of image processing, computer vision and natural language processing, all of that is a very complex procedure which can be done by deep learning models such as convolution neural network and Recurrent neural networks a hybrid model of them will do the desired task [3]. In this paper, the deep neural models have been used to describe an image completely capturing all its details and fully understanding the action verb performed in the image.

Although seeing such a long list of challenges, the model has shown good results. The methodology used for the proposed model is quite effective and has shown some good results with the combination of both models, it was yielding better results. The high level view, as shown in Fig. 2a, is the image vector of size 2048 with the partial caption to feed-forward network which predicts the caption it actually tells the probability of a particular word, the model picks that word which has highest probability and add to the existing caption, the loop moves on till it captures the caption for the whole image and then return the captions of the input image [4]. These methods will give grammatically correct sentences. The model training is very important, the type of dataset, its size and variety in the dataset helps to retrieve



(a) High level Architecture



(b) Caption generation loop

Fig. 2 **a** High level architecture. **b** Caption generation loop

more captions [2]. In the Fig. 2b, shows the caption generation works with the help of deep learning models as well as neural networks for generation of the caption the caption starts "startseq" it totally depends on the designing to use whatever keyword, in the image as it is predicted "the" which may have the highest probability among other vocab words which were used by our model. As soon as we encounter "endseq" out loop will break and the caption will be retrieved.

The work is inspired by the encoder–decoder model, which makes our deep learning algorithm work together as a hybrid to retrieve desired result. Although training the desired model was bit difficult as it was complicated to train model at the same time and create the dataset for the same

2 CNN-RNN-Based Framework

People explore various things around and then classify on the basis of colours mostly, but computer uses slightly different approach to do this [5]. It works with pixels and work on that for them an image is nothing but just a 2D or a 3D matrix of numbers for making them understand it we need to flatten it and then feed them, but the problem was how to extract all the necessary features from an image [6]. In this case we have CNN to do the task for extracting of all the desired images from an image and then using it for further pre-processing but for such a large dataset how do we get all the important features as the dataset is large enough to extract the important features. What if we can use the pre-trained models which were already trained by others and used their weights to do our task. We can use Resntet-50 for the same as it is giving highest accuracy till now in the imagenet competition and we can easily access it from keras library and download the weights.

As shown in the Fig. 3 the model of our Resnet-50 it produces the desired output in the last layer but to extract image features which is of size of 2048 and do that for every image is the requirement in this dataset. This process of work of image feature extraction the model has been trained with the help of captions as show in Fig. 2a. In this propose work, the image captioning task is based on encoder–decoder model, the encoder part is a CNN model for extracting image features which is done by this process. There are various models which can be used such as AlexNet, VGG and GoogleNet. In the second part of the of the model it uses the word vector to perform that it used pre-trained model for RNN which is google glove embeddings and feed it to the RNN model. For each word, it is first represented by, and then goes through the embedding model, this is how our encoder–decoder works together.

The proposed model can give long and complex text as description of the image. It used RNN to store the previous weights and keep continuing forwards for the same it used LONG SHORT TERM MEMORY it is actual an advance version of existing RNN and much better. In Fig. 4 the flowchart of our CNN-LSTM layers

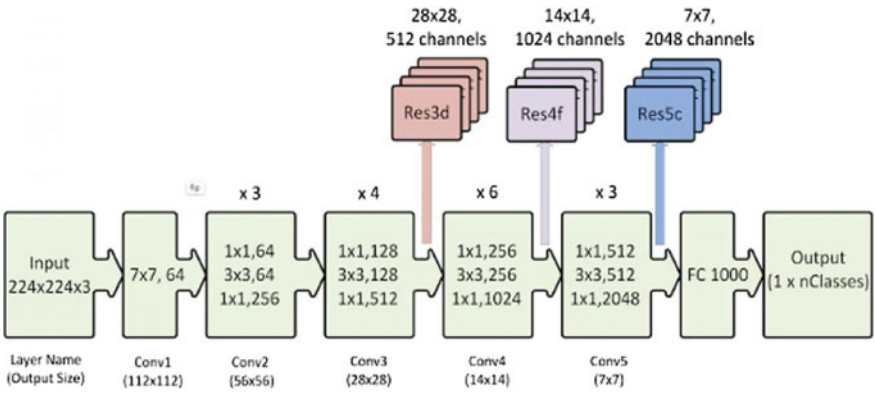


Fig. 3 Resnet-50 architecture

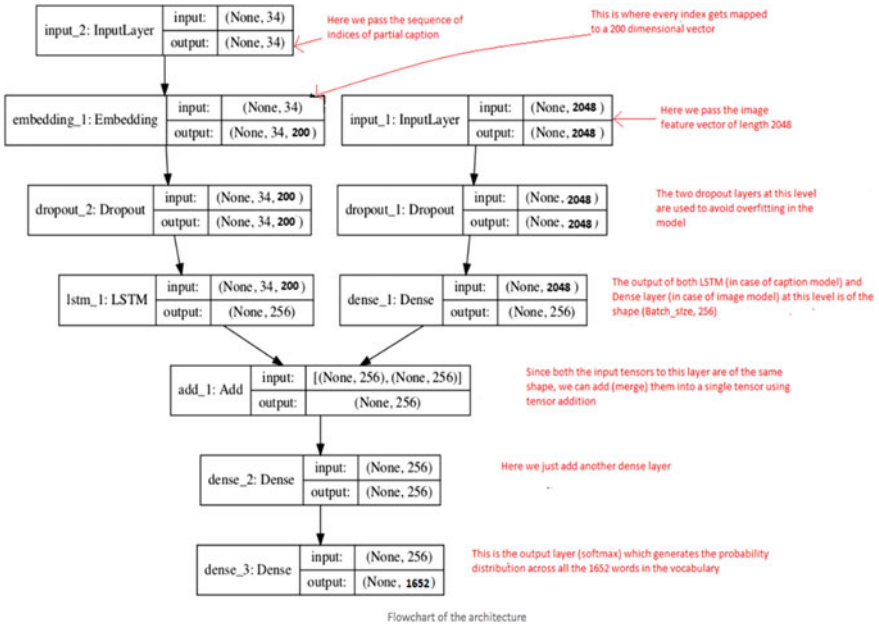


Fig. 4 Flowchart

working together and giving the probability distribution of the vocabulary which was created in the previous stage. Initially the model results in the image vector and integer vector for words. For getting different image description texts the model used different datasets to train the model for getting the desired result for that it used generators which combines both the image and the word together and creates the database. Many datasets are available which will help in training this type of model one of them is FLICKR8K/30 K this dataset divides the training and testing images with a txt file for showing all the captions of an image. Five different captions are given for every image with different length and written in completely different way which helps in describing an image in different ways and with good vocabulary.

The main problem in this framework is about the loss of metrics which is not an optimize solution, as there are numerous ways of describing an image and all can be correct so how to identify the best method among all. The accuracy parameter differentiates between the models for better significant results to prevent the loss.

3 CNN-CNN-Based Framework

Unlike LSTM (Long short term memory) models known as a kind of RNN, which remembers selective sequence of patterns for a long time, better than RNN, but it is still refreshed at each time, which reduce the drawn out memory rather troublesome.

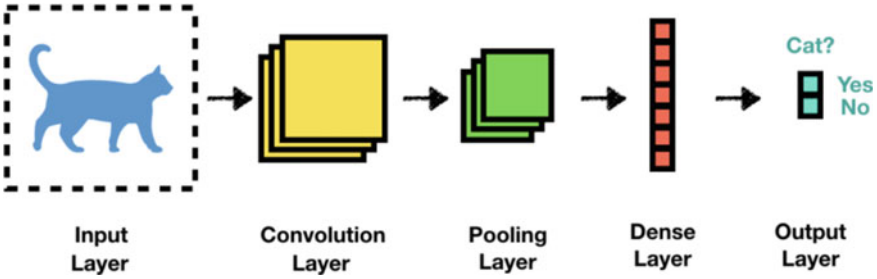


Fig. 5 The process flow of CNN model

The continuous development in the field of Artificial Intelligence, propelled the advantages of CNN models in detecting, analysing and describing images. CNN model capabilities are better utilized by Natural Processing architects. In the field of neural machine interpretation, it has demonstrated that the CNN convolution model is utilized to override the RNN repetitive model, delivering more precised results hence speeding up by a factor of nine. The CNN structure contains three principle segments like the RNN strategy. The first and the last parts are word saving in the array of string containing all the captions. In any use case of dataset, while the middle segment contains LSTM or GRU (Gated Recurrent Unit) units in the RNN case, it covered convolutions which are utilized in the CNN-based methodology. This segment, in contrast to RNN, is feed-forward with no intermediate results saving. Some of the earlier researchers explored CNN-CNN models claiming quick response in fetching all the number of parameters. The accuracy of results was found better in CNN than RNN [5]. The proposed model at times picks the image object but not the frequency of the objects present in the image. As, shown in Fig. 5 the layered deliberation of convolution and the triple door of repeat assume the normal task of image feature identification.

Along these lines, as far as precision, there isn't a lot of contrast between convolutional model and repetitive model. In any case, the way that CNN is quicker than RNN preparing which is straightforward and uncontroversial. The unavoidable outcome is influenced by two parameters [7].

- Convolutions can be run in parallel leading to faster response as compared to RNN which works serially.
- The GPU chip can be utilized to accelerate the preparation of the convolution In the field of machine interpretation and image captioning.

4 Discussions

4.1 Human-Computer Interaction

With the headways of science and innovation and the requirement for the advancement of human life, robots have been utilized in an ever increasing number of businesses. Auto-pilot robots can keenly keep away from obstructions, switch to another lane and people on foot dependent out and about conditions as per the encompassing driving condition they watch. Despite protected and proficient driving, it is additionally conceivable to perform activities, for example, programmed stopping. Freeing the driver's eyes and hands can incredibly encourage individuals' lives and diminish wellbeing mishaps. On the off chance that the machine needs to accomplish the work better, it must cooperate with people better. The machine can mention to people what it sees, and people at that point perform relevant handling depending on machine's response. To finish these undertakings, we have to depend on programmed age of image depictions.

4.2 Image and Video Annotation

In the situation when a client transfers an image, the image should be represented and clarified which can be effectively found by different clients. The customary strategy is to recover the most comparative picture in the database for explanation, yet this technique frequently brings about erroneously commented on images. Plus, video has now become a key piece of individuals' lives. So as to appreciate motion pictures or GIF's better, numerous motion pictures currently require captions. Consistently, there are an enormous number of recordings created around the world. These recordings are made out of a huge number of pictures. In this manner, picture and video comment are an overwhelming undertaking. The programmed age of the image depiction can process all the video casings, and afterwards naturally create the comparing content portrayal as per the substance of the video outline, which can enormously decrease the outstanding task at hand of the video labourer and can finish the video explanation work proficiently and adequately. Moreover, picture and video comment can likewise assist outwardly debilitated individuals with understanding an enormous number of recordings and pictures on the Internet.

5 Conclusion and Future Scope of Work

The CNN models have been proposed in the year 2017. There were numerous enhancements utilizing this structure in machine learning techniques and neural networks which can likewise be applied in image captioning. The CNN-CNN-based

system is a hybrid model between CNN and RNN. Due to their viability in PC vision, CNN ends up used in wide application and much work has been carried out in machine learning techniques. In this paper these enhancements of deep neural networks can be applied in the subtitling of images in the convolutionary model. The image description can be widely used in intuitive inspection, human-machine inter- connection, picture and video comments. This is just piece of the picture subtitling applications. To put it plainly, image captioning can in reality be applied in numerous parts of individuals' lives, which can enormously improve work proficiency and encourage individuals' life, creation and learning. This work can be extended by adding voice to the image description thus helping the visual impaired people for better aesthetic and hedonic experience in life.

References

1. Xin, J. (2017). *Image recognition method based on deep learning*. IEEE.
2. Wettscherreck, D., Aha, D. W., & Mohri, T. (1997). *A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms.*
3. He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep residual learning for mage recognition.*
4. Williams, N., Zander, S., & Armitage, G. (2006). A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification. *ACM SIGCOMM Computer Communication Review, 36*(5), 5–16.
5. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). Going deeper with convolutions. In *CVPR*.
6. Montufar, G., Pascanu, R., Cho, K., & Bengio, Y. (2014). On the number of linear regions of deep neural networks. In *NIPS*.
7. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2014). *Imagenet large scale visual recognition challenge.* arXiv:1409.0575.

Comparative Re-evaluation of Different Single Image Defogging Techniques: A Review

Asifa Baba, Md. Imtiyaz Anwar, A. H. Moon, and Arun Khosla

Abstract Image defogging has turned out to be a demanding task nowadays and to propose several assumptions and approaches for the visibility enhancement of the imagery under consideration, is a need for an hour. Poor weather conditions degrade the image contrast resulting in image blurring and pixel distortion and are thus responsible for the road accidents occurring in the world. For several other applications like remote sensing, video surveillance, navigation, etc., clear and high-quality images are needed. The main objective of the paper is to contemplate the existing state-of-the-art techniques for image defogging for the visibility improvement. Finally, the paper is concluded with the current status of image defogging techniques, their comparative results, and delivers insightful discussions and prospects for future work to boost the efficiency and accurateness of existing systems.

Keywords Computer vision · Image fusion · Fog removal · Performance evaluation · Deep learning

1 Introduction

In most of the image processing applications, limited visibility in foggy environment has been a severe concern. Consequently, under foggy circumstances, the imagery

A. Baba (✉) · A. H. Moon
Islamic University of Science and Technology, Awantipora, Kashmir, India
e-mail: asifa.baba@islamicuniversity.edu.in

A. H. Moon
e-mail: ayaz.moon@islamicuniversity.edu.in

Md. Imtiyaz Anwar
Babasaheb Bhimrao Ambedkar University, Bihar, India
e-mail: imtiyaz.ece@gmail.com

A. Khosla
NIT, Jalandhar, Punjab, India
e-mail: khoslaak@nitj.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_10

acquired by the camera loses their contrast and colour fidelity. In foggy environments, there exist numerous atmospheric particles of considerable size. These particles not only take in and disperse the reflected light of the scene but also disseminate a portion of atmospheric light to the camera. Therefore, the image captured has poor visibility and decreased contrast influencing badly the visible light optical system. This leads to the difficulty in detection of objects under observation and leads to increase in accidents on road. The degradation in visibility is mostly affected by fog during winters in north India and other similar regions of the world. It has been estimated that more than two million people die per year of car accidents across the world, out of which 24% of all such accidents are due to bad weather [1]. In December 2018, eight people including seven from the same family died of road accident in a highway pile-up due to heavy fog in Haryana, India. Immediately on December 29, 2018, seven people were killed and four injured due to heavy fog on Ambala-Chandigarh National Highway, Haryana, India. As per the data given by the Ministry of Road Transport and Highways (MORTH) [2], the number of accidents in India due to foggy weather in years 2016, 2017 and 2018 increased drastically and is shown by the bar graph (Fig. 1).

Therefore, there is a need for an efficient image defogging algorithm to increase the perceptibility and the contrast of the captured degraded image. The images captured are of different types like visible, infrared, computed tomography (CT) images and magnetic resonance images (MRI). In order to defog the image captured by the camera, several visible image defogging techniques are implemented by different researchers. The evaluation of such existing state-of-the-art techniques is contemplated in this paper.

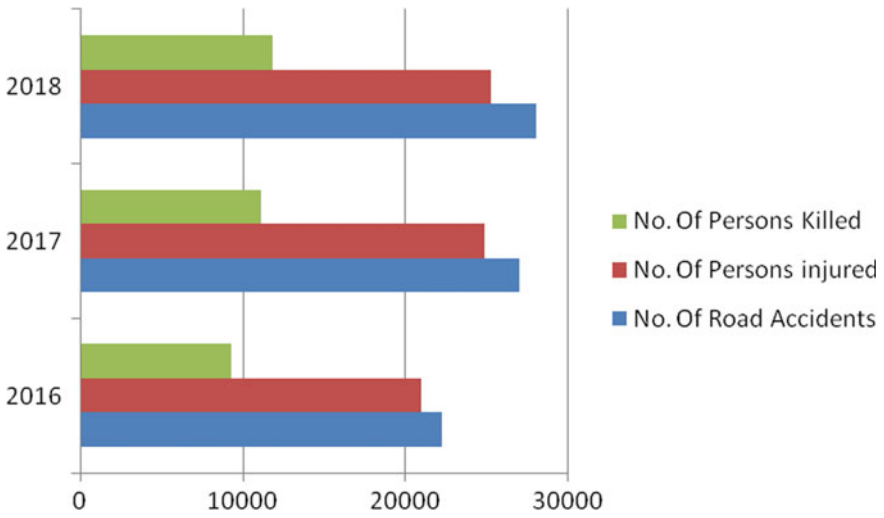


Fig. 1 Number of accidents due to foggy weather in India from 2016 to 2018

2 Review of Contemporary Techniques

The significant work has been done by the researchers in the area of image defogging in order to increase the contrast and colour fidelity of imagery, but the methods in use still struggle with several limitations and improving visibility is an inevitable task to be done. In this section, various visual-, infrared- and fusion-based image defogging techniques have been comprehensively investigated, and their limitations are discussed.

Most of the non-sky patches of haze-free outdoor images include pixels having very low intensities in at least one colour channel. Subsequently, a dark channel prior method has been employed by the researchers in [3] for effective dehazing of a single image to remove the low-intensity information elements. In the proposed technique, the image is more refined by the process of matting in order to reduce the blocking artifacts. The only disadvantage with this algorithm is that the defogging scheme proved to be inefficient for the images containing large sky regions because the colour of the sky is very similar to the atmospheric light in the hazy image.

A visual image defogging scheme has been put forward in [4] in which the visibility of the captured scene in an intense foggy state has been improved. The proposed method has been accomplished by using the methods of contrast enhancement operation, luminance weight map and chromatic weight map followed by the multiscale fusion to get the fog-free images. In order to test the proposed algorithm, large set of images is created by the authors named as SAMEER-TU database to study the effect of variability of emerged scenes. However, under decreased illumination conditions, foggy images still suffer from low visibility.

Authors in [5] have set forth a novel single image fog removal technique for greyscale and RGB images, introducing dark channel prior (DCP) followed by weighted least square (WLS) and high dynamic range (HDR) for the improvement of the visibility of an image in terms of contrast adjustment and preservation of edges without employing any pre-processing steps. However, the proposed technique due to the image conversion to HDR shows minor saturation in few images. Moreover, said proposed scheme is used for the enhancement of the images containing the small sky regions only.

In [6], authors have presented a novel method of fusing visible and infrared images using latent low-rank representation (LatLRR) for visibility enhancement. In this proposed technique, both infrared and visible images are decomposed into two scale representations to generate the low-rank parts and the saliency parts for each source image using LatLRR. In order to reduce the image artefacts, a guided filter is then used on the saliency part, thereby making complete use of spatial reliability. Later, a fusion global-local-topology particle swarm optimization (FGLT-PSO) is used to construct adaptive weights of the low-rank parts in order to get more information from visible and infrared images. The resultant image is recovered back by summing up the low-rank and the fused saliency parts of an image. The said proposed method outperforms several existing fusion methods; however, it is slow in speed and takes much execution time with the increase in the number of iterations.

A novel technique to enhance the contrast of foggy images using contrast limited adaptive histogram equalization (CLAHE) algorithm has been presented in [7]. CLAHE disintegrates the whole image into well-defined regions irrespective of taking the whole image into consideration as in histogram equalization, by determining the number of histograms corresponding to each data region. In this defogging scheme, the RGB image on the one hand is decomposed into red, green and blue components and a CLAHE algorithm is applied to R component of the RGB image and then the CLAHE transformed R component is merged with G and B component of the RGB channel. On the other hand, the same source image is decomposed into HSV components and to S and V components and CLAHE algorithm is applied which are finally merged with the H component of an image. The two resultant images are fused together to get the enhanced image. The proposed system proved to be superior than existing state-of-the-art methods like histogram equalization, adaptive histogram equalization, etc, however the method employed is slower than existing schemes as the CLAHE operates on small distinct image regions rather than the whole image.

An adaptive hybrid image defogging (AHID) algorithm has been presented by the author in [8] for the enhancement of foggy images employing the use of multilevel fusion based on weight maps after enhancing the lighter and darker regions independently. The dark channel prior (DCP) algorithm is employed to approximate the amount of fog present in the image and the fusion-based defogging algorithm is used to remove the remaining image artefacts present in the foggy image. Later, the edges of the resultant image are enhanced using unsharp masking for improved results. Fog removal by the proposed technique is beneficial compared to other conventional methods only when the nearby foggy images are considered; however, the far off object to be enhanced introduces some blurring effects and colouring artefacts in the reconstructed image.

A novel fast algorithm for defogging the single image has been presented in [9] where an RGB image is transformed into HSI form which is later on divided into a plurality of blocks. The maximum point from the S component in each block is selected, and at the same time, I component of HSI image is adjusting so that the fog component is estimated through bilinear interpolation. The fog component thus obtained can be subtracted from the RGB values of each block of an image to be defogged, and the resultant image is enhanced by adjusting the brightness of an image to get the defogged image. The proposed algorithm is simple, easy to implement and showed improved image visibility compared to conventional algorithms. Moreover, the sky regions are dealt individually for the enhancement purpose. However, since the defogging operation is implemented on block level, the technique is slow and takes much execution time.

Employing dark channel prior (DCP) methods for the visibility improvement has a drawback of overestimating the haze from its actual value in case of images containing bright white objects, leads to the distorted image at the output. So, In order to get the reconstructed image with decreased colour distortion researchers in [10] has come up with the image defogging model integrating saliency detection with dark channel prior. The technique of saliency detection method is based on

the contrast of superpixel intensities and is used to get the saliency feature out of the foggy image. These saliency features of the dark channel image, excluded from bright white objects, are used to approximate the transmission and atmospheric light. In order to recover the scene radiance, self-adaptive upper bound is fixed to avoid few regions to be too bright. Since author has used the DCP at its initial stage therefore the proposed algorithm is simple and effective. However, the approach employed in [10] is not used to recover the degraded colour components in the image and hence shows a slight distortion in the reconstructed image.

Authors in [11] employ the technique of depth estimation, colour analysis and a visibility restoration for improving the poor visibility of images collected in calamitous weather. The satisfactory results of median filter technique and an adaptive gamma correction method have been exploited in the proposed depth estimation module for the restoration of images with complicated structure. The transmission map estimation can be adjusted to decrease the halo effects present in the captured image. The colour analysis module has been used to analyse the coloured features of the raw hazy image followed by the visibility restoration block that uses both the transmission map and the colour correlation information to repair the distortion in colour at the end. This technique is used to reduce the atmospheric effects seen in the foggy image with complicated structures, and hence, proficient transmission map estimation can be estimated. The limitation of using median filtering is that the images having small signal-to-noise ratio produce false noise edges by breaking down the edges. Moreover, median filtering cannot curb Gaussian noise distributions.

In [12], a novel image restoration method based on contrast limited adaptive histogram equalization (CLAHE) and no-black pixel constraint with planar assumption (NBPC + PA) methods has been proposed for visibility enhancement. By employing the cascade scheme, the limitations of CLAHE like distorted edges, degradation in colour and halo effects can be abridged. The two approaches together can enhance the visibility and hence offer better results for homogenous as well as inhomogeneous fog. However, due to these cascaded stages, colour saturation has been seen in the output image and hence leads to image artefacts.

In order to obtain the images with high visibility, authors in [13] proposed a restoration technique based on fusion strategy for defogging. Two-derived images have been obtained by the contrast-based method and statistical-based approach, respectively, and are weighted by a specific weight map in order to restore the image back. Compared to other conventional methods, the fusion strategy-based method discussed above proved to be simpler and shows slightly better results. The only limitation of such fusion-based methods is the objectionable visible image artefacts observed due to contrast-based restoration methods employed in the proposed technique.

In [14], a technique of image defogging has been put forth by the authors in which a multilevel perceptron has been used to find out the transmission map directly from the image. Further, dynamic range has also been increased by different contrast stretching methods in order to increase the SSIM index and hence the peak signal-to-noise ratio. However, since the proposed technique is trained on 80 different images, it increases the computational time for the execution of code.

A simpler yet effective method of image defogging has been presented by the authors in [15] involving the process of Retinex algorithm and wavelet transform method, thereby enhancing the boundary information and the high-frequency details of a foggy image. However, the discrete wavelet transform introduces the blocking artefacts in the reconstructed image.

Organization of the paper is as follows. In Sect. 2, review of various contemporary image defogging techniques and their limitations are discussed. Section 3 shows their comparative study in tabular form. Section 4 shows the qualitative and quantitative results of some of the designed algorithms discussed in literature survey, followed by the last section which elucidates the conclusion and future scope of the image defogging techniques.

The contribution of the research conducted by various authors with the limitations found so far in the field of image defogging has been tabulated below (Table 1).

3 Comparative Results and Analysis

In this section, few methods given by the researchers discussed in the literature survey have been evaluated on several test images, and their quality parameters have been calculated and compared for the qualitative and quantitative appraisal. Figure 2 depicts the resultant defogged image obtained using algorithms designed in [3, 14, 15] by the authors, and Fig. 3a, b, c shows the quality parameters like standard deviation, SSIM index and entropy of the defogged images.

4 Conclusion and Future Scope

In this paper, several image defogging techniques for visibility improvement, put forth by researchers in order to evaluate the performance, have been summarized. Despite the considerable progress achieved in such defogging techniques as discussed earlier, a number of limitations have been observed in one way or the other. There also exist several kinds of environmental variations that make the existing techniques more difficult to implement. Therefore, an efficient algorithm is compulsory to muddle through several challenges arising from the nature of visibility enhancement of foggy images. Further, strategy to investigate the research to be carried out includes the use of fusion-based methods for increasing the visibility of the foggy image due to the number of complimentary characteristics of visual and thermal images. There is an enormous scope in this field for further enhancements of foggy images by incorporating the deep learning artificial neural networks especially convolution neural networks (CNN) because of its advantages like maximum utilization of unstructured data, elimination of unnecessary cost and ability to produce high-quality results. Moreover, the performance of different existing techniques may be

Table 1 Summary of comparative findings

Author	Defogging techniques employed	Outcome	Limitations
He et al. [3]	Proposed a simple yet effective single image haze removal technique using dark channel prior followed by the soft matting procedure to refine the transmission by reducing blocking artefacts in an image	• Recovery of high-quality haze-free image containing non-sky regions • Easy algorithm to put into operation	• Proved inefficient for the images containing sky regions
Huang et al. [11]	Presented the technique for visibility improvement utilizing the technique of depth estimation, colour analysis and visibility restoration	• Decrease in halo effects observed in foggy image • Effective map transmission can be achieved • Reduction in colour distortion	• Results in edge breaking • Cannot curb Gaussian noise distributions
Pal et al. [4]	Designed a novel image defogging scheme for the enhancement of dense foggy imagery using SAMEER-TU database	• Easy and simple algorithm used hence incorporates less computing transparency	• Fails to improve the contrast of foggy images captured in a very low illumination state
Thulasika et al. [13]	An image fusion-based strategy has been employed based on contrast enhancement and statistical approach	• Easy to implement due to simple algorithm • Provide better PSNR compared to other state-of-the-art techniques	• Unpleasant image artefacts can be seen in contrast-based restoration methods
Wang et al. [9]	The proposed defogging technique is based on the colour space of HSI. Fragmentation and bilinear interpolation has been employed for effective image defogging	• Improved image visibility than conventional methods • Deals with the sky regions individually • Algorithm is easy and simple	• Because of block level operation, the proposed technique is slow and takes much execution time

(continued)

Table 1 (continued)

Author	Defogging techniques employed	Outcome	Limitations
Anwar et al. [5]	Employs fog removal technique based on dark channel prior, weighted least square and high dynamic range	• Proposed technique proved efficient in preserving sharp details • Maintains the colour quality of the reconstructed image	• Due to the conversion from RGB to HDR, oversaturation in few images has been seen • Suitable for defogging the images with small sky regions
Surekha and Naveen Kumar [15]	Authors presented a novel image defogging technique involving the process of Retinex algorithm and wavelet transform method, thereby enhancing the boundary information and the high-frequency details of a foggy image	• Simpler algorithm used to implement the proposed defogging technique	• DWT leads to artefacts in the output image
Salazar et al. [14]	An image defogging has been presented using multilevel perceptron, thereby increasing the PSNR and SSIM index	• Provides superior performance in terms of SSIM index and PSNR	• More computational time required
Zhang et al. [10]	Designed an image dehazing algorithm by combining saliency features with dark channel prior to get the fog-free image with less colour distortion	• Simple and effective algorithm	• Slight colour distortion in defogged image • Degraded components in an image cannot be improved
Kumar et al. [7]	Proposed image defogging technique based on contrast limited adaptive histogram equalization (CLAHE) algorithm	• Proved superior than existing state-of-the-art methods	• Produced colour and edge distortion, boosts noise and creates halo effects • Slower than existing schemes as it operates on small data regions

(continued)

Table 1 (continued)

Author	Defogging techniques employed	Outcome	Limitations
Han et al. [6]	Put forth a fusion-based adaptive multi-defogging method for improving the visibility	• Better performance than other existing fusion-based methods in terms of evaluation parameters	• Execution time increases with the increase in number of iterations
Hassan et al. [12]	Put forth a cascaded technique employing contrast limited adaptive histogram equalization (CLAHE) and no-black pixel constraint with planar assumption (NBPC + PA) for improving the visibility of foggy image	• Improved performance in terms of visibility enhancement	• Image artefacts observed in restored image
Krishnan et al. 2019 [8]	Presented an adaptive hybrid image defogging (AHID) algorithm using multilevel fusion technique	• Efficient in removing the image artefacts • Better quality of defogging image in terms of PSNR	• Blurring effects and colouring artefacts seen for the far off objects in the foggy image

conducted and compared with existing systems, thus avoiding the chances of road accidents due to bad weather and hence increases the visibility.

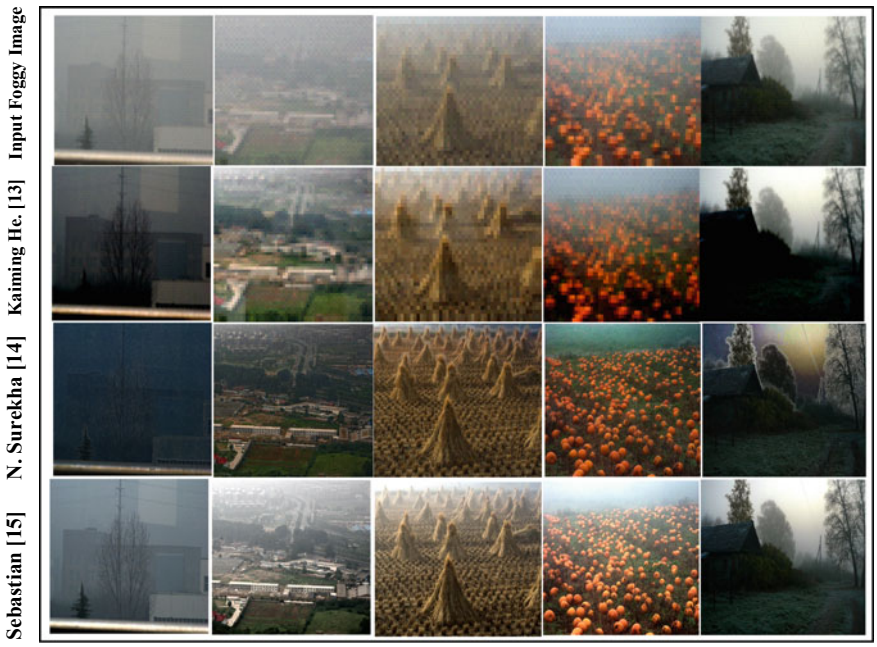


Fig. 2 Performance of defogging algorithms presented in [3, 14, 15] on five different images. From left to right: Scene1, Canon, Cones, Pumpkin, Scene2

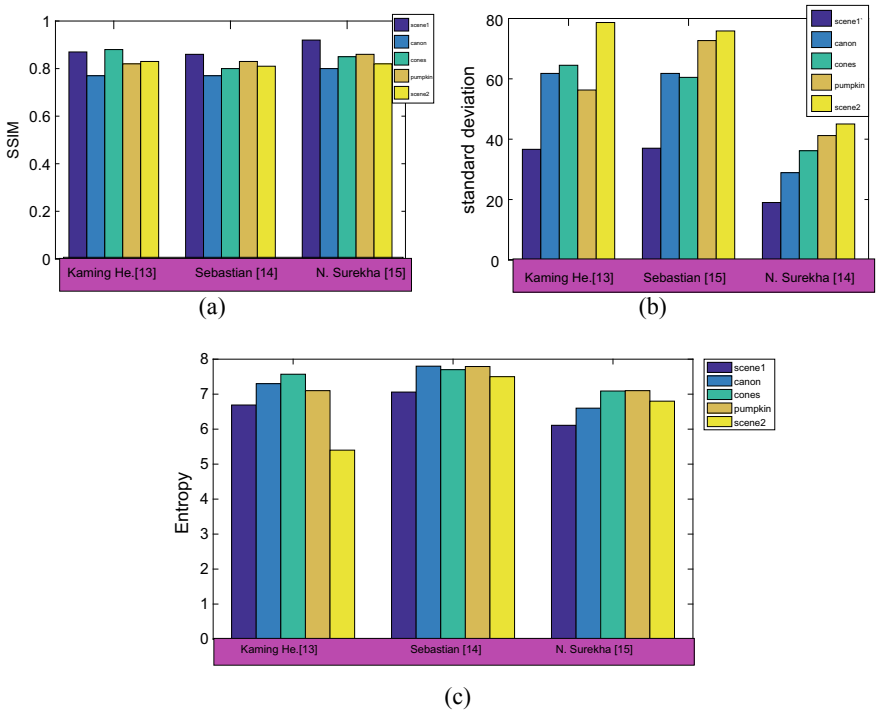


Fig. 3 **a**, **b**, **c** Comparisons of the value of standard deviation, SSIM index and entropy on several test images by the methods employed in [3, 14, 15]

Acknowledgements This research work was supported by Technical Education Quality Improvement Programme (TEQIP-III, under MHRD, Government of India) as a part of CRS Project titled: Development of Fusion-based Defogging Technique for visibility improvement.

References

1. Ding, M., & Wei, L. (2015). Single image Haze removal using the mean vector 12-norm of rgb image sample window. *Optik International Journal for Light and Electron Optics, 126*(23), 3522–3528.
2. MORTH Homepage. https://morth.nic.in/sites/default/files/Road_Accidednt.pdf. Last accessed February 2, 2020.
3. He, K., Sun, J., & Tang, X. (2011). Single image Haze removal using dark channel prior. *IEEE transactions on Pattern Analysis and Machine Intelligence, 33*(12), 2341–2353.
4. Pal, T., Bhowmik, M. K., & Ghosh, A. J. (2014). Defogging of visual images using SAMEER-TU database. In *International Conference on Information and Communication Technologies* (Vol. 46, pp. 1676–1683). Elsevier.
5. Anwar, M. I., & Khosla, A. (2017). Vision enhancement through single image fog removal. *Engineering Science and Technology, an International Journal., 20*(3), 1075–1083.

6. Han, X., Lv, T., et al. (2019). An adaptive two scale image fusion of visible and infrared images. *IEEE Access, 7,* 56341–56352.
7. Kumar, M., & Jindal, S. R. (2018). Fusion of RGB and HSV colour space for foggy image quality enhancement. *Multimedia Tools and Applications, 78*(8), 9791–9799.
8. Krishnan, S., Sabarish, B. A., et al. (2019). Adaptive hybrid image defogging for enhancing foggy images. *Journal of Engineering Science and Technology, 14*(6), 3679–3690.
9. Wang, X., Guob, S., Wang, H., et al. (2016). A fast algorithm for Image defogging. *Proceedings of SPIE, 9684*, 9684271–9684277.
10. Zhang, L., Wang, S., & Wang, X. (2018). Saliency-based dark channel prior model for single image haze removal. *IET Image Processing, 12*(6), 1049–1055.
11. Huang, S.-C., Chen, B-Ho, & Wang, W.-J. (2013). Visibility restoration of single Hazy Images captured in Real World Weather Conditions. *IEEE Transactions on Circuits and Systems for Video Technology, 24*(10), 1814–1824.
12. Hassan, N., Ullah, S., Bhatti, N., Mahmood, H., & Zia, M. (2020). A cascaded approach for image defogging based on physical and enhancement models. In *Signal, Image and Video Processing*. Springer.
13. Thulasika, V., & Ramanan, A. (2015). Single image fog removal based on fusion strategy. In *Fifth International Conference on Advances in Computing and Information Technology* (pp. 115–123).
14. Salazar, S., & Cruz, I. (2018). Single image dehazing using multilayer perceptron. *Journal of Electronic Imaging, 27*(4), 043022.
15. Surekha, N., & Naveen Kumar, J. (2016). An improved fog-removing method for the traffic monitoring image. *International Journal & Magazine of Engineering, Technology, Management and Research, 3*(9), 2061–2065.

A Framework for Implementation of Learning Analytics and Educational Data Mining in Traditional Learning Environment

Anagha Vaidya and Jatinderkumar R. Saini

Abstract The use of the learning analytics (LA) and educational data mining (EDM) is going on from the last two decades. LA is a human-lead process which predicts learners' performance as well as identifying potential problematic issues of the learner as the EDM translates the data into meaningful actions to support and empower the learning. In literature, these practices are used for online learning platform. In face-to-face teaching, these practices are rarely applied. The academic analytic systems are used for collecting the student, teacher and course data, but unable to create actionable and perceptive reports to the educational stakeholders. This paper designed a model for implementation of LA and EDM practices in face-to-face teaching environment. The model used the data about student's continuous evaluation marks, the question paper or learning events used for the assessment. The different processes are generating the reports from it. The system is implemented by R. Shiny.

Keywords Learning analytics (LA) · Educational data mining (EDM) · Learning dashboard (LD) · Academic analytics · Predictive analysis · Prescriptive analysis

1 Introduction

The quality of higher education is considered good, and goals are achieved if education contributes to sustainable employment, personal development and active citizenship [9, 18]. This objective drives the change in the educational model. Two decades back, educational model was more focused on the teacher while now the trend has changed, and it has become more student-focused. Therefore, educational quality measurement parameters are changed. It is focused on the student

A. Vaidya (✉) · J. R. Saini
Symbiosis Institute of Computer Studies and Research, Symbiosis International Deemed University, Pune, India
e-mail: anaghavv@gmail.com

J. R. Saini
e-mail: saini_expert@yahoo.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_11

learning process rather than end-term results. The output of this measurement is the changes in the course design structure, prediction of how many students' retention how much in the course, design of the recommended system, support for personalized learning, etc. Student's learning measurement systematically measured through learning analytics (LA) and educational data mining (EDM) by using technological support like learning dashboard (LD) [8, 17]. The student learning process directly measures the teacher's feedback continuously.

These measurements are accomplished through LA and EDM technology by collecting the data from the student's assessment, student interaction with the courses, student solving the assignment, etc. The assessment identifies what extent of learning objectives have been achieved and provides information to students about their knowledge, skills they are expected to possess after completion of the course [10].

In literature, implementation of LA and EDM practices are used in e-learning, MOOC learning environment in which students interact with the system and their behavior, their learning style, the feedback of teaching material, etc. can be evaluated very easily [12], and the collected data is analyzed for predicting student performance, learning pattern, abnormal fall in performance, needed changes in teaching approach, etc. [15, 17].

In traditional, teaching-learning environment (TLE) is teacher-focused. The teachers are evaluating the students on the end-term written exam papers which are focused on student's memory skills and writing skills only. But the student assessments must include student's learning measurements, quality of the work and the student's efforts for learning [10, 13].

Nowadays, most of the universities have introduced a continuous evaluation process to overcome the limitation of end-term evaluation. In continuous evaluation process, the students are evaluated on different skills like a case study, presentation, quizzes, etc., but still these evaluations are not foolproof, and it depends upon the teacher's expatriation or experience. A systematic assessment model is required, this paper designing a model of the utilization of learning analytics and educational data mining in the TLE. The outcomes of the model are different actionable reports which are used by different educational stakeholders for taking further action.

2 Literature Review

Learning analytics (LA), educational data mining (EDM) and academic analytics are three facets of the existing educational model for the teaching and learning process [6]. The research on LA and EDM is evolving from last two decades, LA is a human-led process which predicts learners' performance as well as identifying potential problematic issues of the learner, [16] LA recommend the students in the selection of new study programs, and it uses the predictive model for actionable analytics for providing the information of individual learners about their progress and the learning risk for higher authority in curriculum development, detections of the complexity of the program, design of learning outcomes, for teachers their teaching skill evaluation,

learning material assessment, etc. [3, 7]. The EDM translates the data into meaningful actions to support and empower the learning [12, 15]. EDM identifies the new patterns from the student's collected data. EDM apply the different data mining algorithm for identification of new pattern from these data. While in LA applies the known models and report actionable report to the system [3].

In education, the different electronic devices are used in teachings like mobile learning, open service platform (MOOC), etc. These devices are generating huge amount of data. The traditional data mining techniques are unable to process theses data; therefore, the new techniques like mobile learning analytics (MLA) and pervasive learning analytics (PLA) are introduced for further analysis [5]. Analyzed data is visualized through learning dashboard (LD) which encourages the learner to engage in more learning activities [14]. The LD displays self-awareness, self-reflection, sense-making and recommendations of the student which help them to improve learning [18].

Student's learning is also measured through emotional, motivation, knowledge understanding and implementation of knowledge in real-life situations [4]. The learning process is strongly connected to learning science [12]. LA is not yet associated with learning science. It identifies the pattern of student behavior but unable to provide underlaying mechanism which needs additional support from learning theories [3]. The learning theory involved student's self-regulated learning (SRL), behavior analysis, motivation measurement, etc. [21] suggested the learning model must relate to learning theories, LA and EDM.

All these concepts are fairly applied to the digital learning environment, but these concepts need to be implemented in face-to-face interaction. This paper tries to develop a model for the same.

3 Proposed Model

The educational environment utilization of computers goes long back. It is used for academic administrative purpose and designed system performs academic analytics [1]. This system is performing descriptive analysis and providing descriptive reports. Some predictive analytics is also performed to an extent and displays the predictive results, but these reports are displayed at the end of the semester which is too late for the decision-makers to take the corrective action against it [2]. Hence, implementation of LA and EDM is required in TLE.

We proposed the model of implementations of LA and EDM practices in TLE. The system is designed by using the layered architecture. The layers are divided into horizontal and vertical layers. The horizontal layers are responsible for processing all teaching and learning processes, and vertical layer is responsible for administrative management like allocation of privileges to the users, log management and third-party connectivity. The model is explained in Fig. 1.

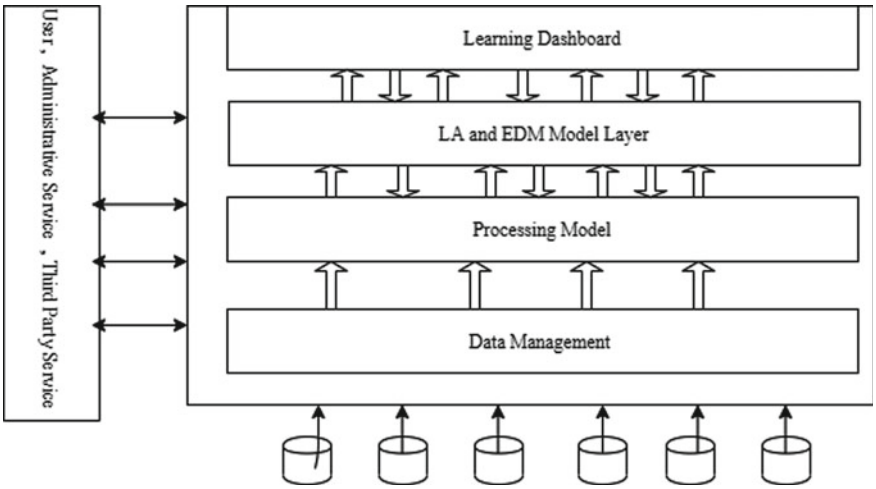


Fig. 1 Block diagram of the system

3.1 Architecture of the System

The input to the model is through student personal information, subject information, learning material, subject planning, subject evaluation planning and attendance report of the students. The subject evaluation planning describes the complete evaluation process of the course. It includes type of assessment and their learning measurement. These learning measurements are quantifying as per Bloom's taxonomy of learning [11].

The functioning of different layers is as follows:

- Learning dashboard layer responsible for answering the different queries.
- LA and EDM model layer: This layer is implementing the LA and EDM models and used different data mining techniques for the implementation. Table 1 displays the different LA and EDM models.
- Process layer: It is designing the services for test generation, student assessment service, course objective analysis, program objective analysis, student learning pattern generation and student behaviors analysis.
- Data management layer: Enormous data is collected through different tests and different evaluation techniques. The input data is also huge. Systematic management is required for this data. This layer provides data services and performs the data cleaning and data encoding activity.

The vertical layer contains the following functionality

A vertical layer consists of managing the user services, security measures and administrative services and providing information from the third-party services (like informing the university rules, etc.).

Table 1 LA models

S. No.	Model	Explanation	Database used
1	Prediction	It performs the prediction of—no. of students drop the course, no. of students will be passed the course, prediction of high and low achievers of the students, predicting next year students will select elective courses, identification of student at risk level (probable failure students)	Student assessment file which measures. The student's domain knowledge and learning activities
2	Recommender system model	Recommendation next semester subject, skill improvement to students and learning activities will be conducted for the subjects	Student assessment and teaching-learning plan
3	Learning measurement	Collaborative learning, self-learning, social learning and learning outcomes of the course. Individual learning outcomes	Student assessment, teaching-learning plan and evaluation plan
3	Student behavior measurement	Identify relationships between a student's poor performance and their behavior during the learning process, students' behavior toward the subject, learning pattern identification, evaluation of learning material. Task complexity, etc., course content evaluation	Student assessment, teaching learning plan and evaluation plan. Student attendance report
4	Teaching feedback	The teaching-learning process, the assessment and learning material, effect of different pedagogical supports provided to learners, how learning material fits the knowledge of the students, student learning performance	Students assessment activity Student attendance report
5	Self-reporting	Student performance analysis, display comparative analysis of result concerning the class	Students assessment activity

Output from system (expectation from stakeholder)

Educational stakeholders are students, teachers, managers and policymakers. Each stakeholders having diverse interests from LA. Table 2 presents the expectation of the stakeholders and how different EDM methods used for evaluation. These expectations are input to our proposed model.

Table 2 Stakeholder expectation

Stakeholders	The expectation from learning analytics	EDM methods used for solution
Student	Semesters report—self-awareness, self-motivation measurement, knowledge measurement	Visualizing the learning progress and achievement to support and behavior measurement
	Need to improve on the skills	Recommendation
	Recommendation for next semester subject	Recommendation
	Next semester progression	Predicative system
Teacher	Feedback about the learning material, learning task assessment	Data mining technique
	To know the student interest knowledge of the student	Data mining technique
	Identification of learners in the class	Learner style identification and learning pattern development
	Evaluation of the test. Identification of anomalous while the evolution of the subject	Anomaly evaluation techniques
Higher management	Teacher-wise report	Predictive system
	Student progression report	Predictive system
	Syllabus revision	Data mining technique
	Course content analysis	Data mining technique
	New course introduction	Data mining technique
	Prediction of student ratio in the next year	Data mining technique
	Prediction of new admission in a course	Predictive system

3.2 Services Used in the System—Student Assessment Service

The process layer implements the different LA practices in the organization. The system architecture section explains the different services for implementation of the cross-section of different service, and their association is mapped into Fig. 2. The inner circle indicates the major actors of the system, and outer circle represents the services provided by the system. From the diagram, 'student assessment' service is one of the important services in the organization. The service is built on continuous evaluation of the students in a semester on the different subjects and on different skills. The time management is the most important factor for this service. The evaluation result should display immediately to the students and teacher, so they can take further action. Therefore, system automation is required, and manual intervention may delay the process. Hence, this model includes an automated assessment system. The service execution is mapped in Fig. 3, and the algorithm is Fig. 4.

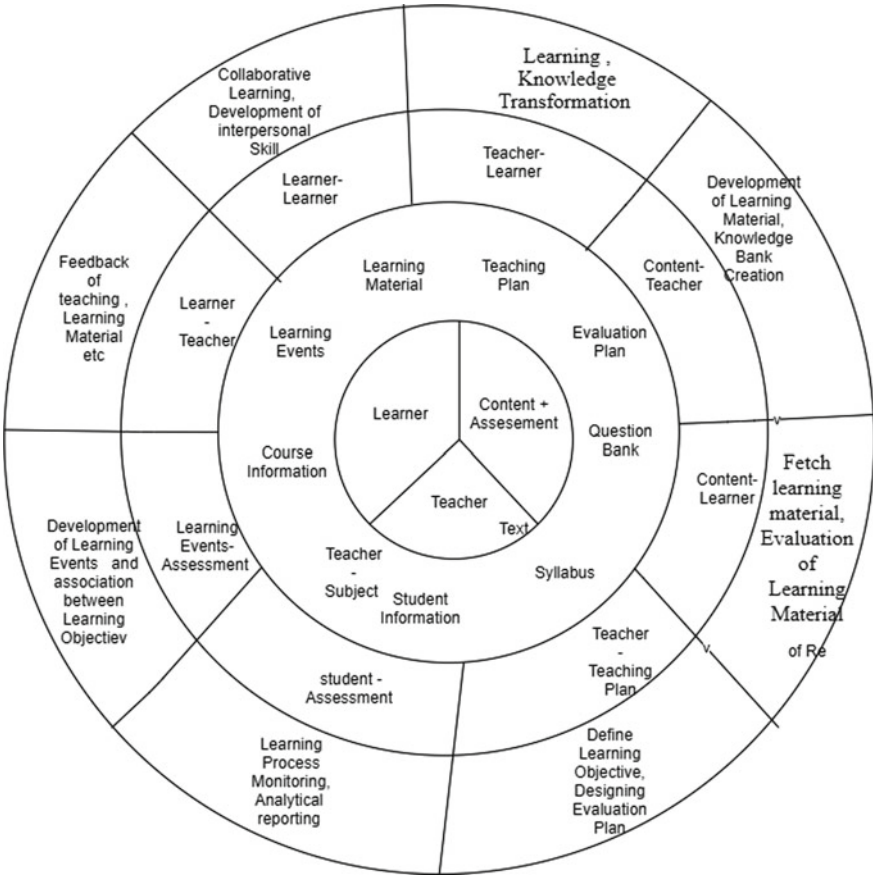


Fig. 2 Cross-section of services in the system

4 Result and Discussion

The framework is developed in R. Shiny interactive web application. The data is collected from a management institute of 360 students. The four-semester data is collected, each semester has average eight subjects, and each subject is evaluated by minimum of three different skills. Instead of creating the database, we are storing the data in excel sheet which reduced the heavy load of the handling the database connectivity and achieve the time complexity and space complexity. The data encoding techniques are developed while storing the data into the sheet. The learning dashboard is developed for reporting the analytical reports to different stakeholder as per their requirements. It displays graphical and textual information of different analyst by using different data mining techniques.

The model represents the predictive and prescriptive reports but needs to put more efforts on graphical representation on this reporting. The line graph and histogram

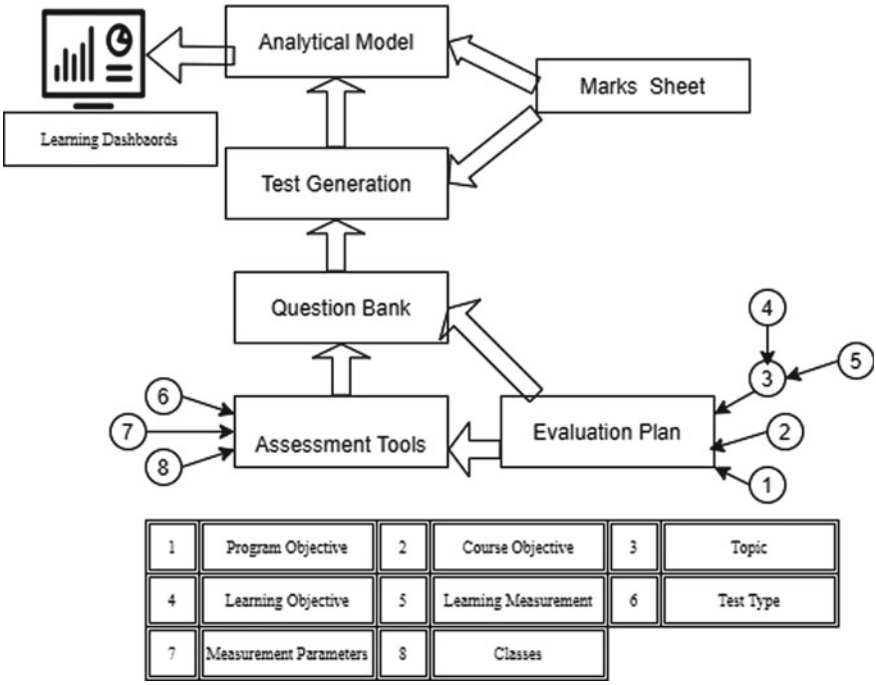


Fig. 3 Student assessment service

Input:- Teaching Learning Plan , Assessment Plan , Question Bank

Procedure: - 1) Read the assessment type from Teaching Learning Plan which contain the information of learning measurement. Read the corresponding assessment parameters and rubrics from assessment plan.

2) In the question bank, the different questions are stored with the attributes of learning skills, blooms measurement level number, measurement of learning skill.

3) As per the assessment type, the question paper is generated from the question Bank along with OMR Sheet.

4) By using from scanner software and R-programming the paper is corrected and question wise marks are stored. In the subjective evaluation the teacher filled the marks on OMR sheets.

5) The evaluation sheets of each student's marks is created

6) The data cube with dimension of learning event wise marks, questionwise-testwise-studentwise marks, overall marksheet of the student is created for the further analysis

Fig. 4 Algorithm for student assessment service

are presented very well with reactive input-output, but it is difficult to include other types of graph. The other problem is on the query optimization. The frequently asked queries need to be organized and query optimization algorithms need to implement for quick response to the stakeholder.

5 Conclusion and Future Work

The paper implements the LA and EDM practice in TLE. The input data is from the existing system, and actionable report is presented to stakeholders. The visual data is presented and makes the decision quicker. It is implemented on computer; in future, it should be operated by using different devices. As number of different devices like mobile, tab etc is used for collection of data and hence needs to think about the common schema for adopting the data management. The implementation of student behavior analysis in future scope.

References

1. Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics, 37,* 13–49.
2. Arora, N., & Saini, J. R. (2014). Predicting student academic performance using fuzzy artmap network. *International Journal of Advances in Engineering Science and Technology,* 3(3), 187–192. ISSN 2319-1120.
3. Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning, 20*(2), 13–29.
4. Bennett, L., & Folley, S. (2019). Four design principles for learner dashboards that support student agency and empowerment. *Journal of Applied Research in Higher Education.*
5. Bodily, R., & Verbert, K. (2017, March). Trends and issues in student-facing learning analytics reporting systems research. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 309–318). ACM.
6. Brouns, F., Zorrilla Pantaleón, M. E., Álvarez Saiz, E. E., Solana-González, P., Cobo Ortega, Á., Rocha Blanco, E. R. … Gabelas Barroso, J. A. (2015). ECO D2. 5 Learning analytics requirements and metrics report.
7. Charleer, S., Moere, A. V., Klerkx, J., Verbert, K., & De Laet, T. (2017). Learning analytics dashboards to support adviser-student dialogue. *IEEE Transactions on Learning Technologies, 11*(3), 389–399.
8. Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2013). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4*(5–6), 318–331.
9. Chomal, V. S., & Saini, J. R. (2013). A study and analysis of paradigm shifts in education triggered by technology. *International Journal of Research in Economics and Social Sciences, 3*(1), 14–28.
10. Dolin, J., Black, P., Harlen, W., & Tiberghien, A. (2018). Exploring relations between formative and summative assessment. In *Transforming assessment* (pp. 53–80). Cham: Springer.
11. Forehand, M. (2010). Bloom's taxonomy. *Emerging Perspectives on Learning, Teaching, and Technology, 41*(4), 47–56.
12. Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018, March). License to evaluate: preparing learning analytics dashboards for educational practice. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 31–40). ACM.
13. Kasat K., Shaikh N., Chandrachud M., & Saini J. R. (2019). Impact of flipped classroom on engagement of post-graduate students under the faculty of social sciences. In *Proceedings of 12th Annual International Conference of Education, Research and Innovation (ICERI-2019)*, Seville, Spain (pp. 7582–7590).

14. Paiva, R., Bittencourt, I. I., Lemos, W., Vinicius, A., & Dermeval, D. (2018, June). Visualizing learning analytics and educational data mining outputs. In *International Conference on Artificial Intelligence in Education* (pp. 251–256). Springer, Cham.
15. Ray, S., & Saeed, M. (2018). Applications of educational data mining and learning analytics tools in handling big data in higher education. In *Applications of big data analytics* (pp. 135–160). Springer, Cham.
16. Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society, 17*(4), 117–132.
17. Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A. … Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*(1), 30–41.
18. Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Assche, F., Parra, G., et al. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing, 18*(6), 1499–1514.
19. Wong, J., Baars, M., de Koning, B. B., van der Zee, T., Davis, D., Khalil, M. … Paas, F. (2019). Educational theories and learning analytics: From data to knowledge. In *Utilizing learning analytics to support study success* (pp. 3–25). Springer, Cham.

Conceptual Framework Model for Opinion Mining for Brands in Facebook Engagements Using Machine Learning Tools

Sneha Rose George, P. Sujith Kumar, and Shine K. George

Abstract The latest pandemic influenza has compelled many business organizations to develop social media marketing strategies. This paper proposes how Facebook can be used by companies for providing brand information and brand interactions at the time of global inflectional disease outbreak. The behavioral outcomes can be measured with the variables such as like, share, and comment. The abundant data which is available on Facebook is impossible to manage manually; therefore, the study also analyses the effective method of handling the big data with the help of machine learning techniques which is an effective approach for sentiment analysis. The paper at hand highlights the key findings and scope for further research.

Keywords Pandemic influenza · Facebook · Sentiment analysis · Social distancing · Behavioral outcome

1 Introduction

The novel corona virus (Covid-19) has profoundly changed the work culture, at least for a period of time. This pandemic influenza which has adversely hit the business forces the brands to engage consumers digitally. As a part of this confrontation, brands have aggravated through brand pages and community interactions, where organizations prefer employees to work from home or give a break to their work life routine and to encourage social distancing. The conventional gathering during fetes, religious rituals, entertainments and other celebrations have come to a phase of pause during this time period. The physical distancing which is inevitable to prevent

S. R. George (✉) · P. Sujith Kumar · S. K. George
Union Christian College, Aluva, Kerala, India
e-mail: sneharosegorge@gmail.com

P. Sujith Kumar
e-mail: sujithparayath@gmail.com

S. K. George
e-mail: shineucc@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_12

the spread of the virus has created all the social community gathering to a standstill, which one way or the other has made the gathering confined to a digital world. The consumption communities came into existence with the introduction of mass media and modern marketing, which is now existing with this virtual presence.

The peer-to peer communication, business communication, and all other means of communication were made easy due to the proliferation of Web 2.0 technology. Social commerce is now a tipping point for the E-commerce, where most of the brands have their social presence. Social networking sites play a crucial role during this scenario by serving as a major pastime for entertainment and gathering information. Among the available social networking sites, Facebook is more accepted among public according to We Are Social (2020) statistics, and there is more than 3.8 bn active social media users worldwide, approximately 93% of these being active mobile social media users.

Understanding their customers is very critical for the brands and market researchers. The company-generated contents and user-generated contents have equal importance in social networking sites mainly in brand fan pages and brand communities. The positive and negative valance creations will directly affect the purchase intention (Hajli 2019). Different tools from netnography to machine learning techniques are available for companies to serve this purpose. To address these challenges, we try to create a conceptual model useful for the brands to understand their customers by analyzing their opinion which is being shared through Facebook.

The remainder of this paper is organized as follows. First, we provide the literature review on need to fill acc to lit review. Second, the conceptual framework and why we have chosen machine learning tools despite of other techniques available. Finally, we provide concluding remarks and suggest directions for future research.

2 Literature Review

2.1 Uses and Gratification Theory

Despite the arrival of many studies on uses and gratifications theory, the first advanced study was introduced in the 1940s by [1] the study of [1] made an effort to describe why users depend on different mass media for deriving different gratifications. And those include television, radio [2] Internet [3], and other social networking sites [4–7]. On the other hand, the introduction of social media technologies revamped the way people use mass media as it is dissimilar in frames and backgrounds. So the way people use social networking sites (SNS) is very much unrelated from the conventional media users who relied on specific media contents [8]. Social networking sites users enjoy the merits of getting information and interesting activities that benefit them. The explanation regarding uses and gratification through Social networking sites has been given by some researchers [4–7]. For raising brand awareness, page

engagement and for increasing communication of the company brands and products, digital marketers are widely incorporating SNSs as an essential key of their online brand strategy. For fulfilling certain needs and wants, people actively seeks out different media is what Uses and Gratification Theory explains [9].

Facebook is one social media that has dominated its growing preference over other social networking sites by the users. For some businesses also it has created dynamics in consumer engagement by promoting their products and services. A user will receive updates when they "like" and "follow" a brand. When the users "like," share, or comment on the post, it would further propagate to their friends and followers. Hence, brand content is transmitted through these Facebook pages at a faster rate and also to a larger audience at a lower cost [10]. The uses and gratifications theory (UGT) which was developed by [11] is been followed in the present study to understand why brand interaction and brand information are important for users which is helpful for brands.

2.2 Relationship Building and Behavioral Outcome

Facebook users will receive information about brands and also they are getting a platform to interact with their favorite brands by doing "like" or "follow" a brand. According to [12], the important drivers for continuous involvement in a brand page is through the likes, comments, and shares. Then, the particular brand followers can "like," share, or comment on the pages. And through a continuous involvement with the contents of the page makes it easier to communicate and also to create a brand reputation within the audience than any traditional media at a less cost [10].

2.3 Gist of Observations

The term "usual business" does not fit these days. The era of social distancing is creating new online behavior among societies as they are not used to sit at home without any social interactions in work life and outside. With the population inside home turned to social media as an outlet that connects family, friends, and work. Social networking sites and issues connected with various crises have been studied, but there are no studies evaluated with a social distancing scenario. Considering the issue, a conceptual framework has been proposed with elements of brand interaction and brand information in regards to uses and gratification at the time of social distancing which has not been studied. For brands at this time of social distancing, social media delivers an opportunity to establish brand reputation by acting as a platform to interact with customers and also to provide a deeper level of information and instant solutions to its customers. Brands are trying to make the interactions with the customers to another level through video interactions, live Q&A sessions, games, and more live time activities. With these interactions, customers are feeling

a close relationship with their favorite brands than ever before. Continued involvement and brand reputation are the relationship outcome of the proposed study and the behavioral outcome being the "like," "share," and "comment."

3 Proposed Conceptual Framework

In this section, the three major parts which are: Uses and gratification during social distancing, relationship building, and behavioral outcomes are been discussed in the conceptual framework. In Fig. 1, the model shows that brand interaction and brand information are antecedents to comment, share, and like in social media. It further states that continued involvement and brand reputation create a bonding between these antecedents and behavioral outcomes. At the time of social distancing, Facebook is the main social networking sites which the organization uses to enhance their brands to consumers. Through brand interaction, there is a continued involvement happening among the customers along with the creation of brand reputation. Brand information on Facebook also leads to generate brand reputation among the followers. Thus, continued involvement and brand reputation bond the user with the respective Facebook page.

The measurement of behavioral outcomes can be done with the "likes," "shares," and "comments" on Facebook. The users will leave a "comment" on the Facebook page when a brand is most respected. Users will "share" the page when the intensity drops to a moderate level and the users "like" when the vigor is dropped from the middle level. Machine learning is used to study the sentiment or opinion mining.

3.1 Why Machine Learning?

Machine learning, the subfield of artificial intelligence, emerged as a way of problem-solving technique. It is used by various businesses as it enhances user experience through social media, which is one of the largest platforms for various marketing purposes like brand creation, business expansion, and promotion. Running a business efficiently on social media is almost impossible without the use of data science

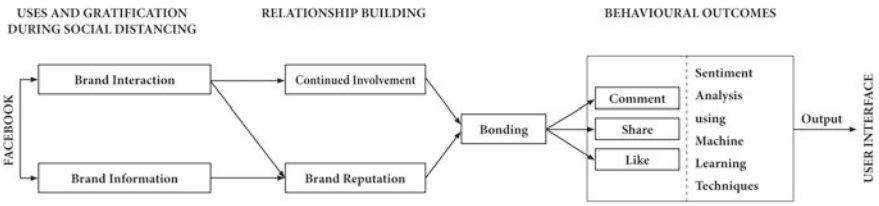


Fig. 1 Conceptual framework during social distancing for opinion mining for brands in Facebook

technologies like machine learning and big data. The purpose of social media which was to provide a connection with friends has changed into serving various brands to advertise and also to expand their customer base. From the likes, comments, and shares, accessing social media profiles made it easy to learn about consumers and consumer behaviors. Big data is helping the organizations to manage the huge data they get from social media for actionable insights. It analyzes a user's social media usage by scrutinizing the posts and pages they like, share, and follow. It thus helps marketers to know what products and services are the individuals likely to be interested in and also in tracking the purchases pattern.

3.1.1 Other Tools Used

The abundant data which is available on Facebook is impossible to manage manually. Tools like netnography are also used to understand the customers as it is very critical for researchers and practitioners. Along with that, other tools like google alerts, google analytics, Facebook insights, tweetstats, people browser, brandwatch, and sematrica'slexalytics are been used by researchers.

3.1.2 Netnography Versus Machine Learning

Netnography is a very useful tool for service research. It is very similar to other qualitative research methodologies for the online environment, such as digital ethnography, online ethnography, virtual ethnography, and cyber-ethnography [13–15], but during the usage of this method, several dilutions are made which does not provide the actual result as more quantification done is considered as one of the main controversies raised on this method. Where in netnograpghy being more into qualifying the digital data, but machine learning is more into the quantification of data which will provide better results. Therefore, an effective method of handling big data with the help of machine learning techniques is an effective approach for sentiment analysis. The positive, negative, and neutral opinions can be mined with the sentiment analysis. Support vector machines (SVMs), Naive-Bayes, and decision trees are predominantly used for sentiment analysis in social media marketing [16, 17].

4 Conclusion and Scope for Further Research

The use of Facebook during the time of social distancing is loud which will be propitious for the brands being digital and frequently interacting. The consumers gradually co-create the contents for the brands through the bonding they have, which leads to electronic word of mouth on Facebook by commenting, sharing, and liking. It will result in positive and negative valance creation through the media. Amidst the

several tools available to measure this opinion mining, we can use machine learning tools which will deliver the essential data for the brands to analyze the sentiments.

Future studies could replicate the conceptual framework to other social networking sites like Instagram, WhatsApp, Twitter, and other more. In addition to this, an empirical study on this model can be used for further validation. More aspects like brand equity, brand loyalty, and customer satisfaction can be included. Furthermore, the sentiment analysis in positive and negative facets can be studied for brands when the situation demands.

References

1. Islam, J. U., Rahman, Z., & Hollebeek, L. D. (2018). Consumer engagement in online brand communities: A solicitation of congruity theory. *Internet Research, 28*(1), 23–45.
2. McLaughlin, C. (2016). Source credibility and consumers' responses to marketer involvement in Facebook brand communities: what causes consumers to engage? *Journal of Interactive Advertising, 16*(2), 101–116.
3. Nisar, T. M., & Whitehead, C. (2016). Brand interactions and social media: Enhancing user loyalty through social networking sites. *Computers in Human Behavior, 62,* 743–753.
4. Ha, Y. W., Kim, J., Libaque-Saenz, C. F., Chang, Y., & Park, M. C. (2015). Use and gratifications of mobile SNSs: Facebook and KakaoTalk in Korea. *Telematics and Informatics, 32*(3), 425–438.
5. Malik, A., Dhir, A., & Nieminen, M. (2016). Uses and gratifications of digital photo sharing on Facebook. *Telematics and Informatics, 33*(1), 129–138.
6. Nyland, R., & Near, C. (2007). Religiosity as a mediating factor in social network use. Paper presented at the midwinter meeting of the Association for Education in Journalism and Mass Communication (AEJMC), Reno, NV, February 23–24. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.168.4424&rep=rep1&type=pdf. Accessed March 2016.
7. Raacke, J., & Bonds-Raacke, J. (2008). MySpace and Facebook: Applying the uses and gratifications theory to exploring friend-networking sites. *Cyberpsychology and Behavior, 11*(2), 169–174.
8. Luarn, P., & Li, H.-H. (2005). Toward an understanding of the behavioural intention to use mobile banking. *Computers in Human Behavior, 21*(6), 873–891.
9. Katz, E., Blumler, J. G., & Gurevitch, M. (1974). Uses and gratifications research. *Public Opinion Quarterly, 37*(4), 509–524. Retrieved April 15, 2012, from: Sciences direct database.
10. Qualman, E. (2013). *Socialnomics: How social media transforms the way we live and do business*. Hoboken, NJ: Wiley.
11. Herzog, H. (1944). What do we really know about day-time serial listeners? In P. F. Lazarsfeld & F. K. N. Stanton (Eds.), *Radio research* (pp. 3–33). New York: Duel, Sloan and Pearce.
12. Cvijikj, P., & Michahelles, F. (2013). Online engagement factors on Facebook brand pages. *Social Network Analysis and Mining, 3*(4), 843–861.
13. Ward, K. (1999). Cyber-ethnography and the emergence of the virtually new community. *Journal of Information Technology, 14*(1), 95–105.
14. Tunçalp, D., & Lê, P. (2014). (Re)locating boundaries: Asystematic review of online ethnography. *Journal of Organizational Ethnography, 3*(1), 59–79.
15. Hjorth, L., Horst, H., Galloway, A., & Bell, G. (2017). *TheRoutledge companion to digital ethnography*. New York, NY: Routledge.

16. Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., & By, T. (2012). Sentiment analysis on social media. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 919–926).
17. Etter, M., Colleoni, E., Illia, L., Meggiorin, K., & D'Eugenio, A. (2018). Measuring organizational legitimacy in social media: Assessing citizens' judgments with sentiment analysis. *Business and Society, 57*(1), 60–97.

Capacity Aware Active Monitoring Load Balancing Principle for Private Cloud

R. Arokia Paul Rajan

Abstract Virtual machines (VMs) are the basic compute elements in cloud computing. There are load balancing principles associated with a job scheduler assigns the requests to these computing elements. Deploying an effective load balancing principle enhances better performance that ultimately achieves users' satisfaction at the high level. Assigning an equal requests load appropriate to the capacity of the VMs will be a fair principle that can be the objective of any load balancing principle. Active monitoring load balancing principle assigns the requests to a server based on the pre-computed threshold limit. This paper presents a technique for assessing the capacity of the VMs based on a common attribute. This work measures each VM's processing ability as a percentage using the statistical method called Z-score. A threshold is quantified and the requests are proportioned based on this threshold value. Each server is then assigned with the proportioned requests. Suitable experiments were conducted Requests Assignment Simulator (RAS), a customized cloud simulator. The results prove that the performance of the proposed principle is comparatively better than a few load balancing principles. Areas of future extension of this work were also identified.

Keywords Cloud computing · Load balancing · Request scheduler · Active monitoring load balancing · Z-score

1 Introduction

A distributed system is a collection of independent computing resources which are the processing elements coordinated with messages. In general, the distributed computing has meant of processing various steps in a business process at the different places by a network of computers [1]. Typically, distributed computing uses the client/server model and it is a widely adapted implementation model. Resource management in a distributed environment is a principle of management of resources

R. Arokia Paul Rajan (✉)
Department of Computer Science, CHRIST (Deemed to be University), Bengaluru, India
e-mail: arokia.rajan@christuniversity.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_13

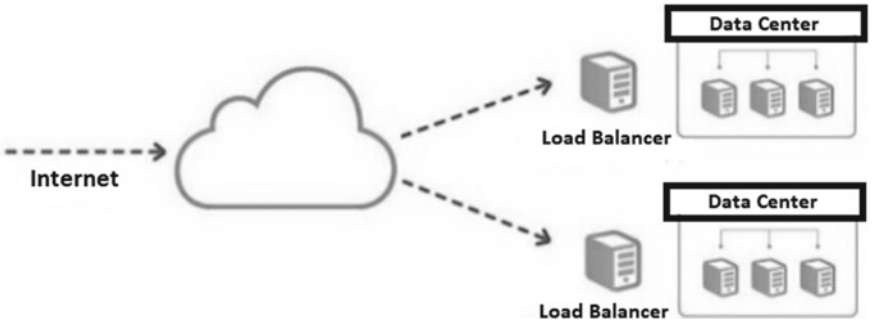


Fig. 1 Load balancing in distributed systems

with the objective of ensuring that the users are accessing the remote resources as fast as he enjoys the local resources. The resource manager is the key component of resource management that controls the assignment of processes with the resources. Resources can be logical entities like shared data, file or physical devices, etc.

Distributed applications which are available in distributed computing environments receive a large number of requests from geographically distributed users and the satisfaction for the consumers is directly proportional to the fastness of their requests service time. A request scheduler is a software component that handles the requests at the background for further execution [2]. Figure 1 presents the resource provisioning through request scheduler in cloud environments.

The load balancing principle determines the best method to assign requests with the servers to achieve the objective of achieving the best performance. Deciding upon the choice of these policies, the users' requests are assigned to a group of computing nodes in the data center [3]. The aim of any load balancing principle is to maximize the resources utilization with minimum response time. The weighted load balancing principle assigns the requests among the set of processing elements considering a static "weight" which is pre-computed [4]. There are principles that weigh the capacity of the servers based on its one or more attributes and splits the incoming requests based on the weight of the each server. Figure 2 presents the illustration of weighted load balancing principle.

This paper is structured as follows: The literature survey discussed in Sect. 2; Sect. 3 discusses the backdrop of the research problem and the proposed algorithm; Sect. 4 discusses the experiments and results; Sect. 5 concludes the work with future research directions.

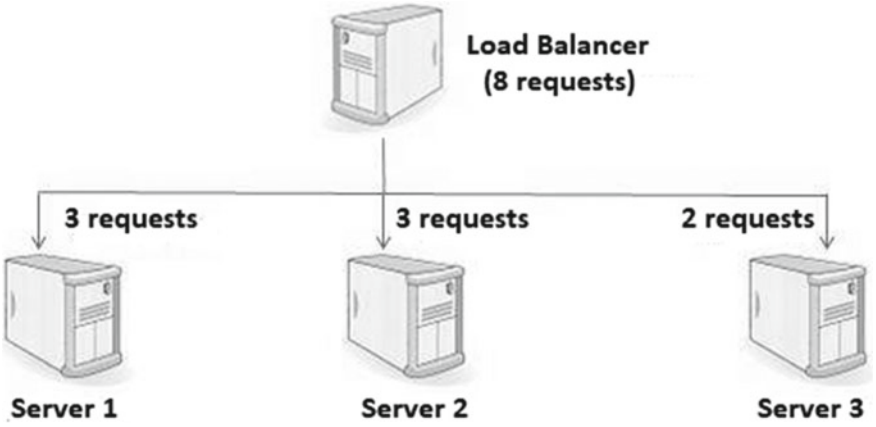


Fig. 2 Weighted load balancing principle

2 Related Works

The significant contributions in the sphere of load balancing principles in the cloud are summarized as follows:

The service scheduling schemes are modelled using a queuing game model which is used in software as a service (SaaS)software as a service (SaaS) [5]. The objective is maximizing the cloud computing platform's (CCP's) payoff via controlling the service requests, whether to join or balk, and controlling the value of the CCP's admission fee. If the position number of a new service request is bigger than the optimal queue length, it balks. Otherwise, it joins in.

Ant colony optimization approach aims efficient distribution of workload among the nodes [6]. When a request is initialized, the ant starts moving toward the source of food from the head node. Regional Load Balancing Node is chosen in cloud computing service provider as a head node. Unprocessed request keeps record of the every node it visited and record their data for future decision making.

A dynamic priority (DP) parallel job scheduler allows users to control their allocated capacity by adjusting their spending over time [7]. This simple mechanism allows the scheduler to make more efficient decisions about which jobs and users prioritize and gives users the tool to optimize and customize their allocations to fit the importance and requirements of their jobs.

The active monitoring load balancer maintains information about each VM's and the number of requests currently allocated to which VM when a request is allocated a new VM arrives. If there is more than one VM, the first identified is selected AMLB returns the VM id to the datacenters controller. The datacenters controller sends the request to the VM identified by that id. The datacenter controller notifies the AMLB to new allocation and cloudlets are sent to it [6]. The weighted active monitoring load balancing algorithm is proposed by modifying the active monitoring load balancer by assigning a weight to each virtual machine.

Using the load balancing algorithm, all VMs are assigned with different amounts of the available processing power [8]. If one VM is capable of having twice as much load as the other, then assigned weight is "2" or if it can take four times load, then assigned weight is "4" and so on.

3 Proposed Algorithm

The premise of the research problem is stated as follows: Consider there is a set of heterogeneous servers with varying capacity based on an attribute. There is a set of requests queued for the servers. The objective is to divide the requests based on the capacity of the servers, with the objective of minimum response time in the queue [9]. Thus, there is a need for a technique which assigns a set of "m" requests to the set of "n" servers. The proposed solution is presented as an algorithm and presented in the below section.

3.1 Algorithm of Capacity Aware Active Monitoring Load Balancing Principle

The primary focus of this research is to devise a technique which will distribute the queued workload across the servers (VMs) evenly. The proposed principle, namely capacity aware load balancing principle aims to distribute the requests proportionate to the servers based on its capacity. There is a need for a technique which will quantify the capacity of the server based on some parameter.

To assess the proportion for each server, this research used statistical method Z-score (or standard score) to quantify each server's capacity based on an attribute's value. Z-score measures a value's relationship to the mean in a group of values of that attribute [10]. Z-Score can be applied in situations to compare a score's relative performance in a group of scores as well as to filter the scores based on a criterion. To illustrate, in a set of 100 students in a class, the performance of a single student in an exam compared with the other students cannot be measured only with mean and standard deviation. In this situations, standard normal distribution and its relative Z-score can be used to measure relative performance. The constraint is the marks of the students' should be normally distributed. Table 1 presents the algorithm which is an enhancement of the existing active monitoring load balancing principle.

This work used number of concurrent connections (load capacity) of the virtual machines as the attribute to measure its capacity. After the enumeration of the request split-off for each server, the requests are assigned from the queue to individual VMs. Also, the requests assignment is subject to the capacity constraints of each VM.

Table 1 Proposed load balancing principle

Procedure Capacity_Aware_Active_Monitoring (m, n)
Input: *m—Number of servers; n—Number of requests*
Output: Map of *m(i)* with *n(j)*

Repeat
1. Let m be the number of VMs.
2. If current allocation count < m, then allocate the VM to the current request;
else create a new list.
3. Let n be the existing request load.
/* Calculate threshold value (number of requests that can be assigned to each server) using Z-score method */
4. Calculate Mean (μ):

$$\mu = \frac{\sum Xi}{n} \quad (1)$$

where X is a set of values of a parameter, n is the number of servers.
5. Calculate standard deviation (σ):

$$\sigma = \sqrt{\frac{Xi-\mu}{n-1}} \quad (2)$$

6. Compute Z-score *(Zi):*

$$Zi = \frac{Xi-\mu}{\sigma} \quad (3)$$

7. Normalize Z-score with normal distribution [11].
8. Enumerate of server's servicing cut-off *(Ti):*

$$Ti = \frac{pi*100}{P} \quad (4)$$

where T is the threshold percentage of each server (Server's servicing cut-off), p is the individual variation of a score and P is the total variation.
9. Calculate Split_off for each server (Si):

$$Si = \left| \frac{R*Ti}{100} \right| \quad (5)$$

where, S—Request split_off, R—Total number of requests in the queue, T—Threshold percentage.
9. Divide the total incoming requests into different sets. This number denotes the server's serving capacity in terms its capacity.
10. Assign the requests to each server.
11. Check whether the VMs are overloaded. If so, distribute the requests to the VMs with least load.
12. Check whether the number of requests are equally loaded for each VM.
13. Compute the response time of the requests.

until all the requests are served;

4 Experiments and Results Discussion

An application has been developed to measure the performance of the devised principle with a few existing load balancing principles. The developed application evolved as a simulator and named as "Request Assignment Simulator (RAS)." The design and parameters of this application are inspired with CloudAnalyst [12]. RAS is a well-customizable, user-friendly visual application that can be able to simulate the process of request scheduling in a distributed architectures. It is designed in such

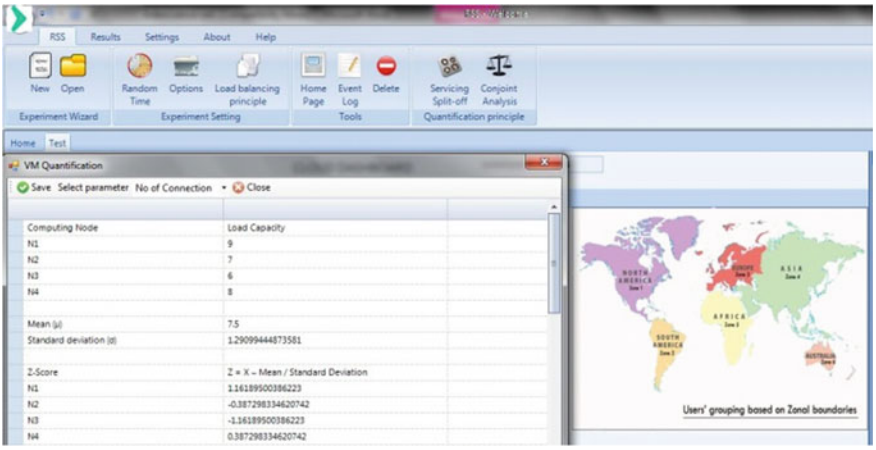


Fig. 3 Experiment using RAS simulator

a way that it can be well fit into Infrastructure as a Service (IaaS) cloud computing model.

Suitable experiments were conducted using RAS and it is observed that the results obtained using the proposed technique showed a considerable improvement in the performance. The devised principle has been compared with round robin and active monitoring load balancing principle.

The wait time, response time, node utilization factor, total earned value by the server, and wait time are a few of the parameters that are observed as the performance parameter. Node utilization factor and total earned value are directly proportional to the performance of the system, whereas average wait time is inversely proportional. A step involved in setting up the experiment in RAS is presented in Fig. 3.

Node Utilization Factor (NUF)

The proposed algorithm quantifies each server's capacity as a threshold number. Based on this servicing cut-off of each server, the requests queue has been split and assigned to each server. The total assigned requests corresponding this threshold number is the actual allotted share proportionate to the capacity of each server. Keeping the server's servicing cut-off values as the threshold for each server, requests are assigned [13]. The other load balancing principles do not have such threshold limit.

NUF represents the number of requests that can be handled by a server proportionate to its capacity. Thus, it is the indicator for the efficient resource utilization. This research work used the server's servicing cut-off threshold as the scaling measure for assessing the performance with other load balancing principles. The comparison of the performances of different load balancing principles based on node utilization factor was observed and presented in Fig. 4.

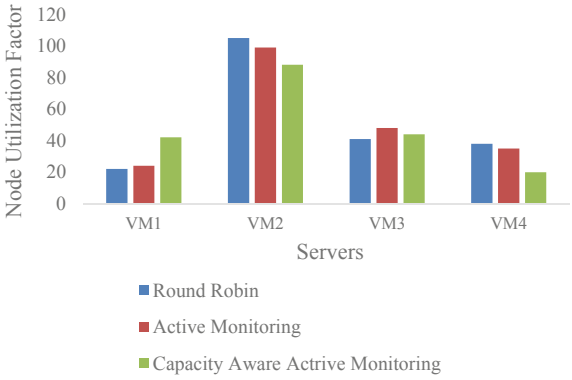


Fig. 4 Comparison of different load balancing principles

NUF is observed by conducting series of experiments and the ANOVA method is performed to analyze the efficiency among the load balancing principles [14]. F-ratio found to be greater than the critical value of the F-table which implies that there is a significant level of differences among the techniques. Therefore, it is evident that the proposed load balancing principle shows a considerable improvement in the NUF compared to round robin and active monitoring load balancing principles.

5 Conclusion

The load balancing principle has a substantial impact on the performance of large-scale distributed computing architectures. Deploying a suitable load balancing principle will ascertain the high performance. There is a need for a principle that will distribute the incoming requests that distributes the load evenly across the servers based on the servers' capacity. This paper addressed this problem and designed a principle which is an extension of active monitoring load balancing principle. This research work designed a suitable method to enumerate each server's allocation share. The statistical method Z-score is used to measure the capacity of the servers. The designed load balancing principle contextually befits Infrastructure as a Service cloud model particularly if it is a private cloud deployment.

Experiments were conducted using a customized application specifically developed for deploying the proposed principle. The experiment results shows an improvement in the performance metrics when it is compared with a few existing load balancing principles. The limitation of the designed load balancing principle is that it can take only the attributes which are numeric data type. Finding a suitable method to assess the capacity of the server using non-numeric attributes will be the future research direction.

References

1. Tanenbaum, A. S., & van Steen, M. (2006). *Distributed systems: Principles and paradigms* (2nd ed.). Prentice-Hall Inc.
2. Marks, E. A., & Lozano, B. (2010). *Executive's guide to cloud computing*. Wiley.
3. Wang, W., & Casale, G. (2013). Evaluating weighted round robin load balancing for cloud web services. In *Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems* (pp. 30–37).
4. Brucker, P. *Scheduling algorithms*. www.users.utu.fi/yurnik/scheduling_files/Scheduling_Peter_Brucker.pdf.
5. Zhang, B., Gao, J., & Ai, J. (2017). Cloud loading balance algorithm. In *International Conference on Information Science and Engineering* (pp. 5001–5004).
6. Rodrigo Calheiros, N., Ranjan, R., Beloglazov, A., De Rose, A. F. C., & Buyya, R. (2011). CloudSim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. In *Software: Practice and experience* (Vol. 41, No. 1, pp. 23–50). Wiley Press.
7. Yi, C., Yi, C., Zhang, X., & Cao, W. (2018). Dynamic weight based load balancing for microservice cluster. In *Proceedings of the 2nd International Conference on Computer Science and Application Engineering* (Article No. 2, pp. 1–7).
8. Wang, W., & Casale, G. (2014). Evaluating weighted round robin load balancing for cloud web services. In *Proceedings of the 16th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing* (pp. 393–400).
9. Nishant, K., Sharma, P., Krishna, V., Gupta, C., Singh, K. P., Nitin, N., & Rastogi, R. (2012). Load balancing of nodes in cloud using ant colony optimization. In *Proceedings of the 14th International Conference on in Computer Modelling and Simulation* (pp. 3–8).
10. Gupta, S. C., & Kapoor, V. K. (2014). *Fundamentals of mathematical statistics*. Sultan Chand & Sons.
11. Panneerselvam, R. (2014). *Research methodology* (2nd ed.). Prentice Hall of India.
12. Wickremasinghe, B., Calheiros, R. N., & Buyya, R. (2010). CloudAnalyst: A CloudSim-Based visual modeller for analysing cloud computing environments and applications. In *Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications* (pp. 446–452).
13. Arokia Paul Rajan, R., & Sagayaraj Francis, F. (2014). Experimenting with Request Assignment Simulator (RAS). *International Journal on Computer Science and Engineering, 6*(11), 363–373.
14. Taha, H. A. (1999). *An introduction to operations research* (6th ed.). Prentice Hall of India Private Limited.

RETRACTED CHAPTER: A Computer Vision-Based Approach for Subspace Clustering and Lagrange Multiplier Optimization in High-Dimensional Data

K. R. Radhika, C. N. Pushpa, J. Thriveni, and K. R. Venugopal

Abstract In this work, we discuss the issues raised due to the high-dimensionality data in real-life scenario and present a novel approach to overcome the high-dimensionality issue. Principal component analysis (PCA)-based dimension reduction and clustering are considered as promising techniques in this field. Due to computational complexities, PCA fails to achieve the desired performance for high-dimensional data, whereas subspace clustering has gained huge attraction from research community due to its nature of handling the high-dimensional data. Here, we present a new approach for subspace clustering for computer vision-based applications. According to the proposed approach, first all subspace clustering problem is formulated which is later converted into an optimization problem. This optimization problem is resolved using a diagonal optimization. Further, we present a Lagrange multiplier-based optimization strategy to reduce the error during reconstruction of low-level data from high-dimensional input data. Proposed approach is validated through experiments where face clustering and motion segmentation experiments are conducted using MATLAB simulation tool. A comparative analysis is presented which shows that the proposed approach achieves better performance when compared with the existing subspace clustering techniques.

Keywords Subspace clustering · Dimension reduction · Computer vision · Optimization

The original version of this chapter was retracted: This chapter was retracted because of a production error resulting in duplicate publication in this book series as different volume. The retraction note to this chapter is available at https://doi.org/10.1007/978-981-15-8354-4_81

K. R. Radhika (✉) · C. N. Pushpa · J. Thriveni · K. R. Venugopal
Department of CSE, UVCE, Bangalore University, Bengaluru, India
e-mail: radhika@bmsit.in

Ś. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_14

1 Introduction

During last decade, demand of data-intensive applications has increased dramatically and an explosive growth is noticed in the data availability. These data are obtained from multiple modalities and sources such as computer vision-based surveillance applications, where multiple cameras are installed in the desired region and different type of data are captured. This process of data collection has led toward various challenges such as data acquisition process, storage and processing the complex high-dimensional data [1].

These types of high-dimensional data-based computations are widely adopted in several areas such as machine learning-based applications, image and signal processing, pattern recognition and bioinformatics. For example, let us consider a computer vision-based application where image consists of huge amount of pixels and videos contains several frames which generate huge data to process for desired task. This process generates high-dimensional data which increase computational complexity and memory requirement of the systems. Moreover, it also affects the performance of applications because of the noise and inaccurate quantity of samples for the given space dimension, known as "curse of dimensionality" [2]. Generally, high-dimensional data are presented in low-dimensional constructions and are not distributed uniformly in the space. Thus, data recovered from the lower-dimension structures can help to reduce the computation complexity, memory requirement and help to recover the performance of pattern inference, machine learning and face recognition tasks. However, any image can consist of large number of pixels, but computer vision models use fewer pixel and considerations to represent the geometry and changing aspects of the given scene. This process has motivated to develop the robust technique which can be used for representing the high-dimensional data into low-dimensional structures [3].

Several studies have been presented to address this issue such as dimension reduction and clustering. In the field of dimension reduction, principal component analysis (PCA) is considered as promising technique where high-dimensional space data are obtained using using low-dimensional space [4]. These techniques are adopted in several applications as discussed before. However, data points can be extracted from numerous subspaces, but the membership subspace data points might be new. For example, in any video sequence, multiple frames are present and multiple moving objects may be presented. Hence, multiple subspaces are required to present these types of scenarios. To overcome these issues, subspace cluster-based techniques have gained huge attraction from research communities. This approach of subspace clustering is used for grouping the data into multiple clusters corresponding to the subspace.

Subspace clustering is a branch of clustering which finds cluster in the subspaces in the given dataset. Generally, these techniques are applied in the high-dimensional dataset where irrelevant and multi-dimensional dataset can result in the noisy clusters. However, feature selection techniques help to remove the irrelevant and redundant dimensions from the data, whereas subspace clustering techniques analyze the data

and identify the relevant data and dimensions and group then into the form of cluster. The subspace clustering techniques are divided into two categories as top-down approach, where initial clusters are formed on the full set and further results are improved in an iterative process; the another approach is bottom approach, where dense regions are identified in the low-dimensional subspaces and later combined to form the clusters.

This approach facilitates the low-dimensional representation from the high-dimensional data by clustering the data into subspace groups of data points. This process is known as subspace clustering and adopted in computer vision applications, for example, image segmentation [5], face recognition using clustering [6], video motion segmentation and image processing [7]. However, several techniques of subspace clustering are present in the literature. If the high-dimensional data are clean where less error or negligible errors are present, then subspace samples can be represented into subspaces easily [8] and the low-dimensional representation can be obtained. If the data suffer from noise and corruptions, then it becomes a challenging task to handle the errors which results in poor low-level representation of the data. Hence, a robust model is needed to handle these types of issues in subspace clustering.

1.1 Issues and Challenges in Subspace Clustering

This section presents a brief discussion about issues and challenges present in the subspace clustering approaches.

- Segmentation and model estimation are strongly coupled to each other. If segmentation data are known, then single subspace can be extracted using PCA [9]. On the other hand, if the subspace factors are known, then the data arguments for finest fit in, respectively subspace also can be identified. During implementation of these schemes, these parameters are not known and hence need to be solved separately for each problem.
- The dissemination of the data inside the subspace is a crucial parameter which is unknown. If the data are distributed around the cluster and cluster center, then centroid clustering-based approaches can be implemented to solve the clustering problem [10]. On the other hand, if data distribution is arbitrary, then central clustering techniques fail to perform the desired task.
- The position of the identified subspaces is an important parameter [11]. If the subspaces are very close and intersecting, then performing accurate subspace clustering becomes tedious task.
- Generally, these high-dimensional data may face issue due to the corrupted data where missing data entries, noise and outlier are the well-known parameters which are responsible for corrupting the data [12].

1.2 Problem Definition

Let us consider that a set of data samples is obtained which may contain different types of errors in linear subspaces. The main aim of subspace clustering is to reduce the errors and rearranging the corresponding segments into their respective subspaces. In this process, the "error" represents the deviation between subspaces and the data which can be presented as noisy data, missing entries, outliers and data corruptions. These techniques can be used in different applications such image processing and data mining video summarization.

1.3 Contribution

In this work, a subspace clustering process for dimension reduction for computer vision-based applications is proposed. First of all, subspace clustering problem is modeled for multiple subspaces and reconstruction of the data is done where noise is present in the original data. In order to perform this task, we formulate an optimization problem which is solved using diagonal matrix-based solution. This problem is further solved using Lagrange multiplier-based scheme until the optimal convergence condition is obtained.

1.4 Organization

The complete paper is arranged as follows: A brief discussion about recent studies is discussed in Sect. 2. The proposed model for solving the subspace clustering problem is discussed in Sect. 3. Sect. 4 presents comparative study, and finally, Sect. 5 briefs about concluding remarks.

2 Literature Survey

This section gives a brief discussion about recent studies in the field of subspace clustering.

Peng et al. [13] discussed the graph-based learning scheme for subspace clustering approach to establish the relation between data points and subspace. However, these techniques suffer from the complexity issues; hence, authors developed a new novel approach to represent the error-free data extracted from the data points. In this work, linear intra-subspace projection dominance is performed to compute the coefficients. Later, a sparsity graph is constructed and denoted as L2-graph.

Shao et al. [14] introduced arbitrarily oriented synchronized cluster (ORSC) approach for processing the high-dimensional data that uses synchronization process to handle the complex processes. Mainly, this approach is based on the weighted communication model and dynamic clustering process. This technique detects the correlation between arbitrary clusters and arbitrary shaped clusters, initially. Later, outlier detection and noise reduction are applied to improve the performance of subspace clustering.

These techniques are widely adopted in real-time computer vision-based applications. Based on this assumption, Javed et al. [15] presented a computer vision-based model for background and foreground detection using subspace clustering technique. However, several approaches have been introduced recently. Due to dynamic background, occlusions and jitters, these techniques fail to achieve the desired performance. To overcome these issues, authors developed a novel approach which uses spatial and temporal process of subspace clustering using robust PCA model. According to this process, motion-aware correlation coefficients are used for computing the spatio-temporal model for the given video sequence and a low-rank model is developed to represent the low-level structures of the video frames.

Tsakiris et al. [16] discussed the subspace clustering on the missing value data that can be very helpful for reconstruction of data from low-level structures. This technique uses sparse subspace clustering (SSC) for low-dimensional representation of the data for pattern recognition and computer vision applications. However, this work gives a theoretical proof that SSC-based techniques can be used for subspace clustering where missing data are present.

Xia et al. [17] focused on human motion segmentation strategy using subspace clustering method. In this work, the segmentation problem is transformed into the temporal subspace clustering problem. Later, sparse subspace clustering is considered, and geodesic exponential kernel is generated to model the low-level arrangement.

3 Proposed Model

This segment explains the proposed approach for subspace clustering. The complete section is divided as follows: First section presents a subspace clustering problem formulation scenario which focuses on obtaining the noise-free low-dimensional structure data from the high-dimensional data. The next phase of the work gives a novel solution for the subspace analysis problem. Finally, implementation scenario using proposed approach for face clustering is explained. Figure 1 shows the complete representation of proposed approach which includes linear subspace creation, aim definition of SSC, decomposition of the subspace clustering problem, optimization solution and performance analysis.

According to the proposed work, first of all we present linear subspace modeling for the given data and the specific aims are designed for the considered databases such as dimension identification, segmentation points and identification of total number

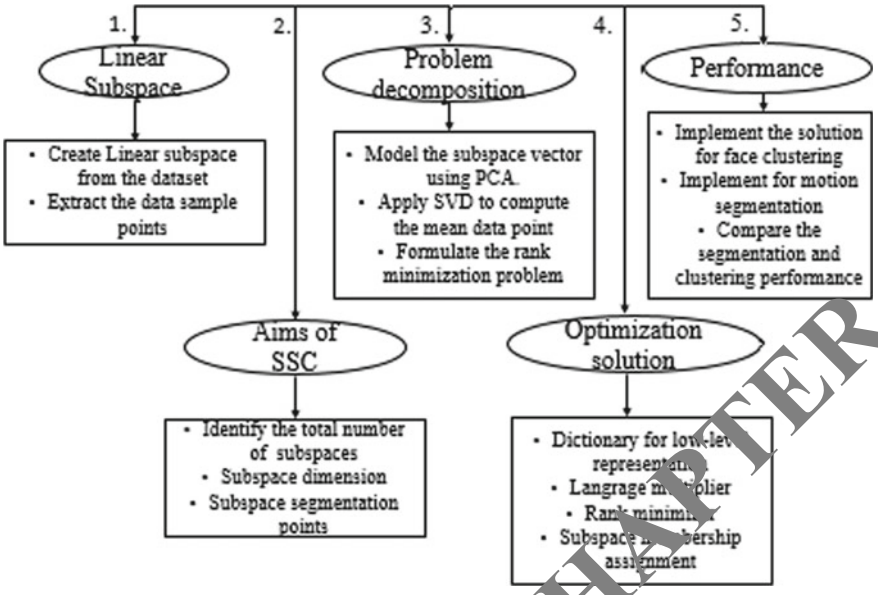


Fig. 1 Proposed work flowchart

of subspaces. In the next phase, problem definition is defined where PCA is used for decomposition of the problem and later SVD is applied to generate the mean data points. This problem is formulated as optimization problem where dictionary-level representation, rank minimization and subspace memberships are assigned. Finally, we apply proposed solution for different types of databases, and the obtained performance is compared with the existing approaches to show the robust performance of proposed model.

3.1 Problem Formulation and Solution

This subsection discusses the problem formulation for subspace clustering.

Let us consider a data modeling problem where multiple data points are present with the combination of subspaces. A sample representation is depicted in Fig. 2, where two lines and one plane make up a three space union as $\mathbb{S}$.

Let us consider that the extracted points are given as x which are extracted from the unknown union of linear subspace with unknown dimensions, i.e., the extracted points are given as $\left\{x_j \in \mathbb{S}^D\right\}_{j=1}^{N}$ that are having linear subspace as $\{\mathcal{L}_i\}_{i=1}^{n}$ along with unknown dimensions as $d_i = \dim(\mathcal{L}_i), 0 < d_i < D$ where $i = 1, 2, \ldots, n$ where $n \geq 1$. Hence, the subspace can be defined as

$$\mathcal{L}_i = \left\{x \in \mathbb{S}^D : x = \alpha_i + \mathcal{U}_i y\right\} \tag{1}$$

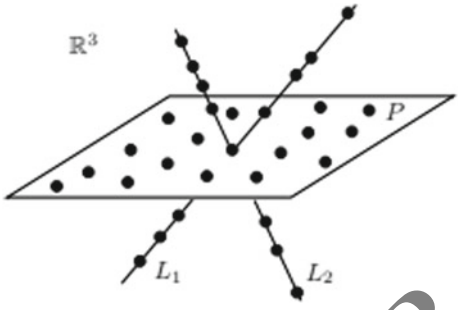


Fig. 2 Sample point set is extracted from $\mathbb{S}^3$ three space union as two lines and one plane

where $\alpha_i \in \mathbb{S}^D$ denotes a random point in the considered subspace $\mathcal{L}$ for linear subspaces $\alpha_i = 0$, $\mathcal{U}_i = \mathbb{R}^{D \times d_i}$ denotes the subspace basis and $y \in \mathbb{S}^d$ denotes the low-dimensional structure representation for the data point x. The main aim of this process of subspace clustering is to estimate the total number of subspaces (n), subspace dimension $\{d_i\}_{i=1}^n$, basis of subspaces $\{\mathcal{U}_i\}_{i=1}^n$, $\{\alpha_i\}_{i=1}^n$ and the segmentation points based on the subspace information.

In this process, if the total number of subspaces is one, then the subspace problem is reduced to computing a vector as $\alpha \in \mathbb{S}^D$, a basis of subspace as $\mathcal{U} \in \mathbb{S}^{D \times d}$ and a low-dimensional representation structure given as $\mathcal{W} = [w_1, \ldots, w_N] \in \mathbb{R}^{D \times d_i}$ and the subspace dimension as d. This problem can be modeled into PCA model which can be resolved using singular valued decompositions (SVD). Here, $\alpha = \frac{1}{N} \sum_{j=1}^{N} x_j$ denotes the mean of the data points which are considered, and the $(\mathcal{U}, \mathcal{W})$ can be achieved using rank d SVD of the mean subtracted data matrix $X = [x_1 - \alpha, x_2 - \alpha, \ldots, x_N - \alpha] \in \mathbb{R}^{D \times N}$, which can be computed as:

$$\mathcal{U} = U \text{ and } \mathcal{W} = \sum \mathcal{V}^{\intercal} \text{ where } \mathcal{W} = U \sum \mathcal{V}^{\intercal} \tag{2}$$

where d can be attained as $d = \text{rank}(X)$ with noise-free data. Let X be a r rank SVD given as $X = U \sum \mathcal{V}^{\intercal}$, $U \in \mathbb{S}^{D \times r}$, $\sum \in \mathbb{S}^{r \times r}$ and $\mathcal{V} \in \mathbb{S}^{N \times r}$. Let us consider that $X_0 \in \mathbb{S}^{D \times N}$ rank SVD values can store a set of n-dimensional sections which are extracted from the subspace union $\{\mathcal{L}_i\}_{i=1}^k$ with unknown dimensions. Hence, observation vector is generated as:

$$X = X_0 + \mathcal{E}_0 \tag{3}$$

Here, our main aim is to improve the row space data of X_0 to improve the system efficiency.

In order to recover the low-level structure data from the observation vector X, we represent this problem as regularized rank minimization problem which can be given as:

$$\min_{P, \mathcal{E}} \text{rank}(D) + \gamma \mathcal{E}_l \text{ subject to } X = D + \mathcal{E} \tag{4}$$

where $\gamma > 0$ and $.l$ denotes the regularization strategy. However, conventional techniques consider that the extracted data are sampled into single subspace; hence, union of subspace clustering may be larger and recovery may be inaccurate. In order to deal with this subject, the problem is redefined as general rank minimization:

$$\begin{array}{cc} \min & \mathcal{M},\mathcal{E} rank(\mathcal{M}) + \gamma \mathcal{E}_l \\ \text{subject to } X = A\mathcal{M} + \gamma \mathcal{E}_l & \text{subject to } X = A\mathcal{M} + \mathcal{E} \end{array} \tag{5}$$

where A denotes the span dictionary of the given data space and $\mathcal{M}$ denotes the minimizer for low-level representation of data X. Once the optimal solutions are identified for the $\mathcal{M}$ and $\mathcal{E}$ as $\mathcal{M}^*$ and $\mathcal{E}^*$, the original data can be recovered using $A\mathcal{M}^*$.

Let us consider that $A \neq 0$ and $X = A\mathcal{M}$ have feasible solution in the given span of data space as $\mathcal{M}^* = A^{\dagger}X$. The minimizer problem can be given as $\text{rank}(\mathcal{M}^*) = \text{rank}(X)$. In this process, dictionary selection plays important role which displays that the proper selection of dictionary can be helpful to obtain the desired low-level representation. This can be achieved when data columns of A and X are sampled exactly and then minimizer problem can help to achieve the subspace membership for the given data points. Here, we denote the collection of subspaces $\{\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_k\}$, where each subspace has a rank $r_i > 0$. We assume that A_i denotes the total m_i samples of the ith subspace $\mathcal{L}_i$, and X_i denotes the sample collection n_i samples. In this case, if the given subspaces are liberated, the optimization problem is denoted in a block diagonal form which is given as

$$\mathcal{M}^* = \begin{bmatrix} \mathcal{M}_1^* & 0 & 0 & 0 \\ 0 & \mathcal{M}_2^* & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathcal{M}_k^* \end{bmatrix} \tag{6}$$

where $\mathcal{M}^*$ denotes $n_i \times n_i$ matrix which contains the coefficients with the given rank as $\text{rank}(\mathcal{M}_1^*) = \text{rank}(X_i)$. This problem is further solved using Lagrange multiplier approach which can be expressed as follows:

$$\begin{aligned} L = \mathcal{H}_* + \gamma \mathcal{E} + tr\big(Y_1^T (X - A\mathcal{M} - \mathcal{E})\big) + tr\big(Y_2^T (\mathcal{M} - \mathcal{H})\big) \\ + \frac{\alpha}{2}\big(X - A\mathcal{M} - \mathcal{E}_F^2 + \mathcal{M} - \mathcal{H}_F^2\big) \end{aligned} \tag{7}$$

The complete process is given below:

Input: Incomplete data matrix as X, $\mathcal{M}, \mathcal{H}, \mathcal{E}, Y_2, = 0, \alpha = 10^{-6}$
Output: reconstructed data matrix, labels and coefficient matrix

Step 1: Initialize the data matrix computation parameters $X_0 \leftarrow initialize$
Step 2: while not convergence
Step 3: estimate the diagonal block
Step 4: fix other parameters and update $\mathcal{H}$ as

$$\mathcal{H} = \arg min \frac{1}{\alpha}\mathcal{H}_* + \frac{1}{2}\mathcal{H} - \left(\mathcal{M} + \frac{Y_2}{\alpha}\right)_F^2$$

Step 5: fix other parameters and update $\mathcal{M}$ as

$$M = \frac{1}{(AA^T+I)}\left((X - \mathcal{E})A^T + \mathcal{H} + \frac{Y_1A^T - Y_2}{\alpha}\right)$$

Step 6: fix other parameters and update error $\mathcal{E}$

$$\mathcal{E} = \arg\min \frac{\gamma}{\alpha}\mathcal{E}_{2,1} + \frac{1}{2}\mathcal{E} - \left(X - A\mathcal{M} + \frac{Y_1}{\alpha}\right)_F^2$$

Step 7: evaluate the convergence stage as
$X - A\mathcal{M} - \mathcal{E}_\infty < \epsilon$ and $\mathcal{M} - \mathcal{H}_\infty < \epsilon$
Step 8: update the diagonal block and rearrange the data according to the obtained updated ranks

4 Results and Discussion

This section presents a complete experimental study using proposed approach for subspace clustering. The complete experimental study is carried out using MATLAB 2013b running on Windows platform. The proposed approach is implemented for face clustering and video motion segmentation purpose. Face clustering is performed using Extended Yale B database [8] and Hopkins 155 database [9]. The proposed approach performance is compared with the several state-of-the-art techniques such as LLR [10], SCC [11], SSC [2], LSRC [1], LSA [12] and ARM [6] in terms of misclassification rate which can be computed as

$$\text{Error} = \frac{\text{wrong clustered points}}{\text{total number of points}} \times 100\% \tag{8}$$

4.1 Face Clustering Comparison

Figure 3 shows a sample outcome of the proposed approach of data reconstruction technique where first column represents the input image, second column represents the error matrix, and the outcome reconstructed image is given in third column.

Based on these experiments, we present a comparative study in terms of clustering error rate and equated the performance of the suggested approach with existing

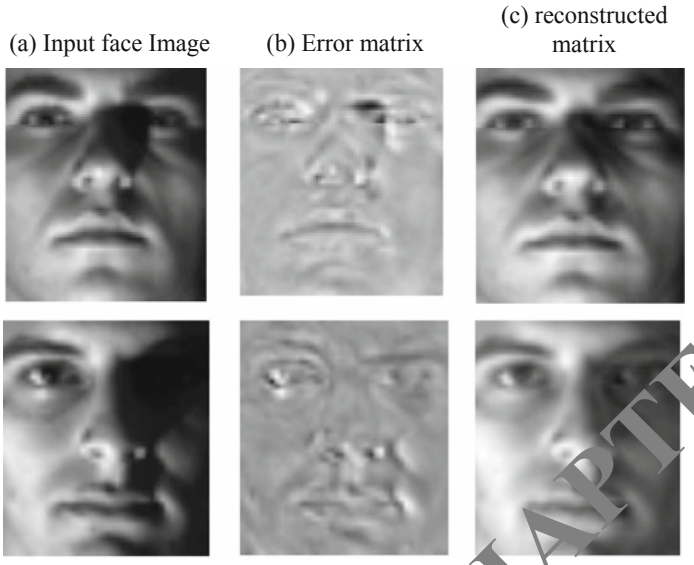


Fig. 3 Sample image, error image and reconstructed image

techniques of face subspace clustering. To show the robust performance, the proposed approach is compared with varied number of subjects and evaluated mean and median of the clustering error rates given in Table

Table 1 Clustering error rate comparis

Algorithm	SSC [2]	LLR [10]	LSRSC [1]	LSA [12]	SCC [11]	ARM [6]	PS
Subject-2							
Median	0.0	0.78	4.69	47.66	7.82	0.78	0.03
Mean	1.56	2.54	5.32	32.80	16.62	1.51	1.42
Subject-3							
Median	1.04	2.60	7.81	50.01	39.05	1.56	0.93
Mean	3.10	4.21	8.47	52.29	38.15	2.26	2.11
Subject 5							
Median	2.51	5.63	12.21	58.08	58.01	3.05	2.11
Mean	4.30	6.90	11.25	59.18	56.81	2.48	2.19
Subject 8							
Median	4.48	10.05	28.03	58.58	59.38	3.32	2.11
Mean	5.84	14.32	23.72	59.18	58.91	3.71	2.59
Subject 10							
Median	5.62	23.58	28.74	57.51	75.77	2.97	2.95
Mean	10.92	22.91	30.31	60.41	73.02	3.85	3.81

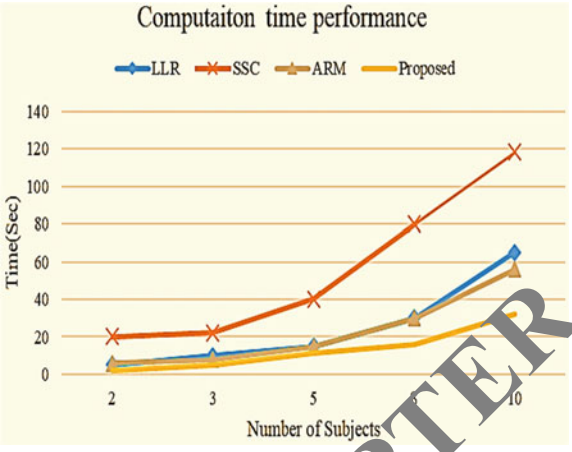


Fig. 4 Computation time performance

Table 1 shows a comparative analysis for varied number of subjects. This study shows that the proposed approach accomplishes better performance when compared with the existing techniques, i.e., for ten subject scenarios, the overall error is 3.81% which is significantly less for existing techniques due to significant cluster formation with the reduced error when compared with existing techniques. Similarly, error variations are also very less in the proposed approach.

Figure 4 shows a comparative performance in terms of computation time using LLR, SSC, ARM and proposed approach which are obtained as 25, 56, 23 and 13.2 s. The proposed approach achieves a desired solution by consuming less time and iteration for convergence.

4.2 Motion Segmentation Performance Analysis

This section presents another experimental study for computer vision application for motion segmentation. This experimental study is carried out using Hopkins 155 datasets, and the performance is evaluated in terms of segmentation error. According to this process of motion segmentation, a video sequence is considered for evaluations and the moving objects are segmented into various spatio-temporal regions which represent a different type of motion.

Figure 5 shows a sample representation of tracing feature during the processing of the input video. This figure represents the background and foreground with different features. The static background maintains similar features during the complete sequence, whereas the features of moving objects vary for each frame. These features can be helpful for various types of computer vision applications such as object detection and tracking crowd behavior analysis.

Fig. 5 Video motion segmentation traced features

Table 2 shows a comparative performance in terms of motion segmentation error. The experimental study for all motion scenario shows that the overall error for all motion is 1.33%, which is comparatively very low in comparison with the state-of-the-art techniques. During the process of data reconstruction, data labeling and coefficient computation, the convergence process also takes place, whereas conventional approaches require more iteration and time to converge, while proposed approach converges fast which in turn helps to reduce the time complexity. Moreover, existing

Table 2 Motion segmentation comparison performance

Algorithm	SSC [2]	LLR [10]	LSRSC [1]	LSA [12]	SCC [11]	ARM [6]	PS
Motion-2							
Median	0.0	0.0	4.69	4.21	0.0	0.0	0.01
Mean	1.51	2.12	5.32	0.55	2.87	1.45	1.40
Motion-3							
Median	0.55	1.42	3.79	1.46	0.55	0.83	0.02
Mean	4.41	4.02	4.85	7.01	4.39	1.48	1.40
All motion							
Median	0.0	0	0.59	0.88	0	0	0
Mean	2.17	2.55	4.85	4.85	4.11	1.47	1.33

approaches suffer from the inappropriate low-rank representation and optimization problem which leads to the poor performance of updating the error matrix. Hence, the proposed approach achieves better results when compared with the existing techniques.

5 Conclusions

A novel methodology for a low-level representation scheme for high-dimensional data using subspace clustering technique is proposed. According to this approach, a subspace clustering problem is formulated for a data where multiple number of subspaces are present and data are corrupted with the noise. Here, optimization problem is addressed by reconstructing the low-level representation of the data, and we formulate an optimization problem which is addressed using a diagonal optimizer solution and Lagrange multiplier scheme to optimize the error parameters. An extensive experimental study is carried out for facial clustering and human motion segmentation. Results of the proposed approach display significant improvement in the performance of subspace clustering.

References

1. Favaro, P., Vidal, R., & Ravichandran, A. (2011). A closed form solution to robust subspace estimation and clustering. In *IEEE Conference on Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 1801–1807).
2. Elhamifar, E., & Vidal, R. (2013). Sparse subspace clustering: Algorithm, theory, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2765–278.
3. Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 395–416.
4. Arias-Castro, E., Lerman, G., & Zhang, T.: Spectral clustering based on local PCA. *The Journal of Machine Learning Research*, 253–309.
5. Yang, A., Wright, J., Ma, Y., & Sastry, S. (2008). Unsupervised segmentation of natural images via lossy data compression. *Computer Vision and Image Understanding, 110*(2), 212–225.
6. Wang, W., Yang, C., Chen, H., Feng, X. (2018). Unified discriminative and coherent semi-supervised subspace clustering. *IEEE Transactions on Image Processing*, 2461–2470.
7. Liu, G., Lin, Z., & Yu, Y. (2010). Robust subspace segmentation by low-rank representation. In *Proceedings of International Conference on Machine Learning* (pp. 663–667).
8. Lee, K. C., Ho, J., & Kriegman, D. (2005). Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 684–698.
9. Tron, R., & Vidal, R. (2007). A benchmark for the comparison of 3-D motion segmentation algorithms. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8).
10. Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., & Ma, Y. (2013). Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 171–184.
11. Chen, G., & Lerman, G. (2009). Spectral curvature clustering (SCC). *International Journal of Computer Vision*, 317–330.

12. Yan, J., & Pollefeys, M. (2006). A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate. In *Computer Vision–ECCV* (pp. 94–106). Springer.
13. Peng, X., Yu, Z., Yi, Z., & Tang, H. (2017). Constructing the L2-graph for robust subspace learning and subspace clustering. *IEEE Transactions on Cybernetics*, 1053–1066.
14. Shao, J., Wang, X., Yang, Q., Plant, C., & Böhm, C. (2017). Synchronization-based scalable subspace clustering of high-dimensional data. In *Knowledge and information systems* (pp. 83–111).
15. Javed, S., Mahmood, A., Bouwmans, T., & Jung, S.K. (2017). Background–foreground modeling based on spatiotemporal sparse subspace clustering. *IEEE Transactions on Image Processing*, 5840–5854.
16. Tsakiris, M. C., & Vidal, R. (2018). Theoretical analysis of sparse subspace clustering with missing entries. arXiv preprint arXiv: 1801.00393.
17. Xia, G., Sun, H., Feng, L., Zhang, G., & Liu, Y. (2018). Human motion segmentation via robust kernel sparse subspace clustering. *IEEE Transactions on Image Processing*, 135–150.

Effective Knowledge Discovery Using Data Mining Algorithm

Garima Sharma and Vikas Tripathi

Abstract In present data science world, data is primary for any analytics, analysis, mining, prediction, and description activity. Although many steps are defined and actively used in random for cleaning and preprocessing of dataset previously but there exist some gaps, degrading the overall quality of data as well as knowledge discovery procedure. In this work, some major additional activities are identified in data manipulation paradigm that can enhance decision-making capability of any mining algorithm. With introduction of new methods at collection and cleaning levels, a data preprocessing algorithm is also proposed here. The improvement in overall knowledge discovery process is demonstrated using a real-time dataset.

Keywords Data quality · Data preprocessing · Data analysis · Data cleaning · Knowledge discovery · Data manipulation · R

1 Introduction

Data preprocessing is a broad umbrella which covers ample amount of strategies and techniques that are correlated and interrelated in many ways [1]. For getting zest of any dataset, a major part lies in its cleaning and manipulation of collected data. As nothing can be perfect, so the same problem lies with our data. Before performing any preprocessing activities like aggregation, dimensionality reduction, or feature extraction, quality of data is cynosure [2, 3]. Any analytics results are solely dependent upon the quality of its training dataset.

In real-world data collection, before moving any data into some statistical algorithm like classification, clustering, mining, etc. [4], there are number of pre-requisite steps to be followed for getting rightful, accurate, and trustful results. Extracting new features from given set of attributes is very common nowadays. Though selecting the right attributes from hundreds of given feature set is a matter of expert, keeping the required ones and eliminating the irrelevant or redundant attributes not only helps

G. Sharma (✉) · V. Tripathi
Graphic Era Deemed to be University, Dehradun, Uttarakhand 248002, India
e-mail: garimavrm91@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_15

in maintaining the data fitness, reducing the dimensionality of data, as well as helps decision-making algorithms to run faster and more efficiently.

This paper proposed an algorithm for effective knowledge discovery by covering more methods for mitigating data quality issues. Our focus area includes incorporation of new steps in data collection and cleaning, which has direct impact on quality of results in knowledge discovery.

This paper is organized in the sections as in Sect. 2, we discussed about steps to be performed at the time of data collection, so that an advance refinement of collected data could be done at the first level itself along with data analysis and cleaning activities, and we have covered few more checks and treatments to be perform in data preprocessing. The proposed algorithm and detailed implementation strategies are explained in Sect. 3, and Sect. 4 showcases the impact of given algorithm in effective and more truthful knowledge discovery. The concluding remarks and future work is given in Sect. 5 showcasing how one can achieve staggering and more reliable results by stepwise implementing the given algorithm.

2 Related Work

2.1 Data Collection

A classical definition of data collection is gathering of information in a systematic fashion. This statement has evolved a lot with time. At present, data collection tools and techniques are way more than just fetching and loading of data. Complete ETL process—extract, transform, load is expected from a collection tool in modern systems. Keeping this into view, we analyzed few datasets and found some abnormalities other than the implemented ones [5]. As there could be more than one source while fetching data into a particular system, a problem of inconsistency exist in column names, i.e., for a similar attribute, there could exist different name from different sources. Second irregularity we found was distinct formats for a single attribute. Most widely seen example for this case is timestamp. Third anomaly proved primer in erroneous result in knowledge discovery phase was incorrect column type, i.e., the attribute data type was not matching with its values. These anomalies are required to be essentially removed at the time of collection and implicitly before analysis.

2.2 Data Analysis and Cleaning

Data analysis is a process of organization of data in drawing helpful conclusions. This phase acts as base for various data cleaning [4] activities. Identification and removal of inconsistent and imprecise values present in any crude dataset is main aim of any data cleaning method. Noise and outlier detection algorithms like clustering or

unsupervised machine learning algorithm work efficiently in searching and removing the abnormal records [6]. Here, abnormal refers to the outliers or oddly present dataset showing serious deviation from other data items present with in the dataset. As unsupervised algorithms has no labels, and therefore, no boundaries exist for framing the data items, thus helpful in finding anomalies. This is classically performed in all the analysis work [7]. Here, we find an improvement in this legacy analysis and cleaning system. As we know, there could exists hundreds, thousands of features in a single dataset [8]; therefore, there may exist a possibility of interconnection and interrelation between them. These relationships can be used in treating missing values as well as NULL values present in the dataset. These relationship values so obtained perform a crucial role in data manipulation. Various machine learning algorithms like apriori algorithm and KNN can be used to predict the missing values using these relationships. Redundancy in records are required to be removed after all data manipulation.

2.3 Data Preprocessing

To get data finally ready for discovering knowledge, it must be passed through data preprocessing phase. This generally includes integration of various attributes and creation of a new attributes by aggregation or segregation of attributes, selecting required and primary features while dropping irrelevant ones [5]. Normalization can be done in the end of preprocessing unit if in case scaling is required. This can be performed in two ways—min-max normalization and z score normalization [9].

3 Proposed Algorithm

Following is the associated algorithm to be executed stepwise for getting maximum data quality and knowledge discovery (Fig. 1):

1. Ingest dataset (K_i) whereas 'i' is the number of sources, 'n' is the number of columns, and 'r' is the number of rows.
2. Check and rectify column name mismatching between similar attributes of different sources and make one unit by combining all the datasets (K).
3. Now, further check number of columns (n) and their formats, i.e., whether all the values are present in a single format or not. If no, correct the same and proceed to step 04
4. Detect the datatype of each column and again check whether they are complementing with respective column values. If no, correct the same and proceed to step 05
5. Start data analysis phase by detecting and removing noise and outliers present in the dataset.

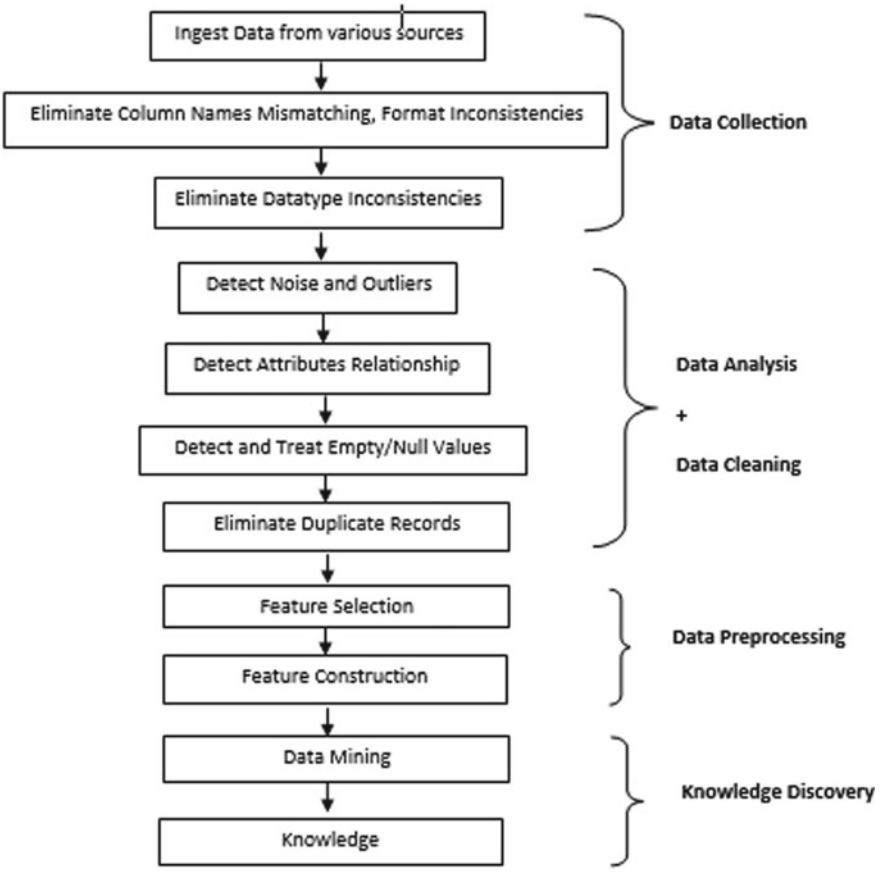


Fig. 1 Proposed data mining algorithm

6. Since there is only one dataset now, detect the relationships between the different attributes. This is useful in data manipulation, i.e., for treating missing and NULL values. Rectify and correct such values using this step.
7. Check duplicate values 'dR' present in the dataset (K), if yes, then go to step 08, else go to step 09.
8. Remove duplicates and check the unique number of rows 'UR'

$$r = \mathrm{dR} + \mathrm{UR}$$

9. Proceed further with other data preprocessing activities like feature selection and construction of new attributes from given attributes by following aggregation, segregation, etc.
10. Finally, knowledge discovery procedure can be begin based upon the use cases.

4 Knowledge Discovery

Knowledge discovery is solely dependent upon the quality of data passing into discovery systems [10]. This involves application of various data preprocessing methods aimed at facilitation of data mining algorithms. Many times even required post-processing for refining and improvement of knowledge [11]. For validation of stated algorithm, we have taken a dataset from NYC open data Web site [12]. The data contains information of dog owners living in New York City. All the residents of New York City are required to license their dogs right after their adoption as per the given law. In our dataset, each record represents a unique dog license issued date and expiry date. Each tuple stands as a unique license period for a dog over a year-long time period. This dataset has 15 columns and 51,861 rows saved in a csv format. After analyzing the dataset attributes, other than problems like null value and missing value, there were major data quality issues, refer Table 1.

To understand, the dirty data here is screenshot of our sample grubby dataset, and its inconsistencies are discussed in Table 1 quality issues column (Fig. 2).

The main aim of this work is to showcase the importance of new methods explored and implemented in data collection and cleaning level. The results shows how the new algorithm is impacting the overall knowledge discovery procedure of given dataset. We have used an open-source statistical language, R, and analyzed the results using exploratory data analysis technique [15].

In Fig. 3, due to presence of imprecise boroughs, the distribution checkup between the NYC states was incorrect, but after finding the relationship between zipcode and

Table 1 Data quality issues and suggested solution

Column(s) name	Messy area	Suggested solution	Quality issue
Animal gender	03 Types—M, F, “ ”	Drop the missing value row	Data inconsistency
Animal birth name	String datatype	Timestamp datatype	Data type inconsistency
Borough	New York City has three different abbreviations Queens has two different abbreviations Staten Islands has two different abbreviations	Single name to one area type	Misspellings
ZipCode	Correct zipcode maps to number of “unknown” boroughs Same zipcode maps to different boroughs	Zipcode–borough mapping correction using correct data [13, 14]	Find relationship between zipcode and borough
License issued date	String datatype	Date datatype	Incorrect datatype
License expired date	String datatype	Date datatype	Incorrect datatype

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
RowNuml	AnimalNar	Animal	AnimalBirthMonth	BreedName	Borough	ZipCode	Commu	CensusT	NTA	CityCo	Congi	State	LicenseIssuedDate	LicenseExpiredDate	
40879	LUCY	F	02-01-2003 00:00	German Shepherd	Queens	11364	411	129104	QN42	23	6	11	11/24/2015	11/17/2016	
52045	BUDDY	M	01-01-2007 00:00	Labrador Retriever	Manhattan	10036	104	129	MN15	3	10	27	02-09-2016	02-09-2017	
52046	BUDDY	M	10-01-2004 00:00	German Shepherd	Staten Islan	10306	503	14606	SI54	51	11	24	02-09-2016	03/30/2017	
52047	BRANDY	F	01-01-2014 00:00	Unknown	Staten Islan	10312	503	17010	SI48	51	11	24	02-09-2016	02-09-2017	
52048	BOZO	M	12-01-2001 00:00	Lhasa Apso	Queens	11373	404	473	QN29	25	6	16	02-09-2016	02-09-2017	
52049	BOOTS	F	01-01-2014 00:00	Unknown	Queens	11416	409	36	QN53	32	7	15	02-09-2016	02-04-2017	
52050	BRUCE	M	09-01-2015 00:00	Rat Terrier	Brooklyn	11228	310	196	BK30	43	11	22	02-09-2016	02-09-2017	
52051	BROOKLYN	M	01-01-2009 00:00	Shih Tzu	Brooklyn	11203	317	946	BK96	45	9	21	02-09-2016	02-09-2017	
52052	LANDRY	M	10-01-2011 00:00	Havanese	Manhattan	10028	108	140	MN40	5	12	28	02-09-2016	11/16/2016	
52053	LIEBE	F	11-01-2006 00:00	German Shepherd	Staten Islan	10312	503	17007	SI48	51	11	24	02-09-2016	03/30/2017	
52054	LACEY	F	11-01-2009 00:00	Unknown	Staten Islan	10314	502	18702	SI05	49	11	24	02-09-2016	02/19/2017	
888	BISCUIT	M	01-01-2003 00:00	Jack Russell Terrier	STATEN IS	10312	503	15603	SI54	51	11	24	11/26/2014	01/30/2016	
1517	MR.PICKLE	M	10-01-2002 00:00	Lhasa Apso	STATEN IS	10310	501	121	SI35	49	11	24	12/23/2014	01/30/2016	

Fig. 2 Screenshot of grubby dataset opened in excel

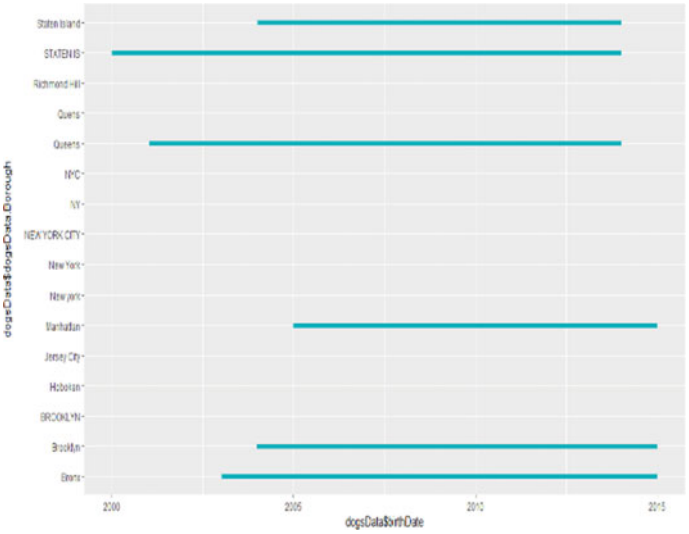


Fig. 3 Number of dogs/town without cleaning

borough (city name), we were able to correct the missing and incorrect boroughs which eventually impact our distribution of dogs in a particular town, refer Fig. 4.

Also, it was not possible to plot a comparison report between birth dates and license issue dates as both the columns were present in string format with different date format types. After correcting the datatypes at data collection time, we were able to set a contrast chart between the two columns, and the results can be seen in Fig. 5.

As number of boroughs were present, it was not feasible to plot a gender distribution chart for all the boroughs individually. After data preprocessing, we were able to set a contrast chart between the boroughs in Fig. 6.

Using Fig. 6, we can even discover information about number of males or females present in a particular borough, we can calculate the ratio between the males and females in a particular town, etc.

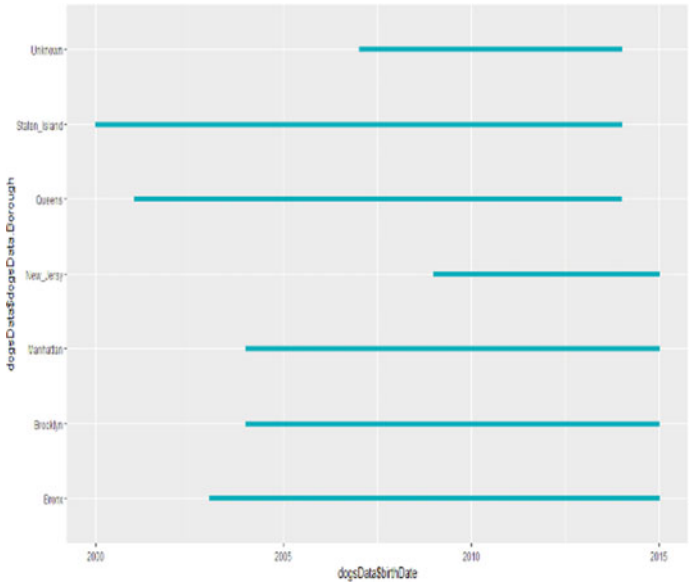


Fig. 4 Number of dogs/town after cleaning

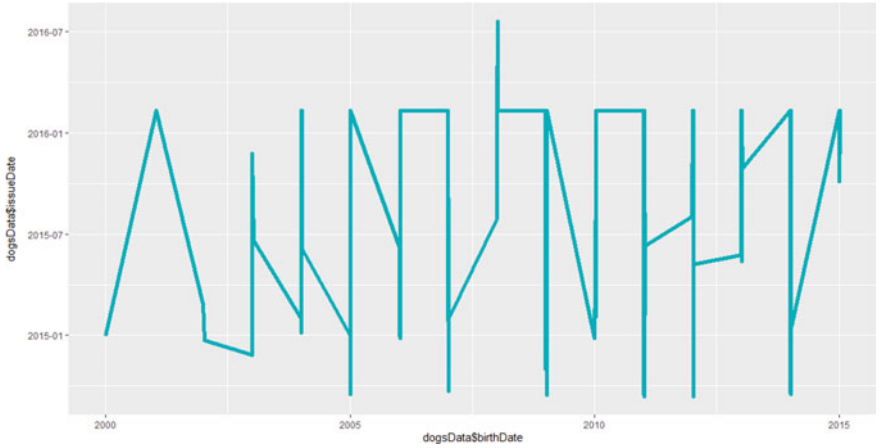


Fig. 5 Dogs birth date versus license issued dates

Lastly, we wished to obtain most favorite breed of New York City, and then again it was not possible with raw data due to presence of enormous number of null values. After treating the null values by setting relationship between the column and dropping all the unknown breeds. After cleaning and preprocessing the transformed dataset, we obtain following word cloud based on the number of counts. Higher the count, more centered the position of value.

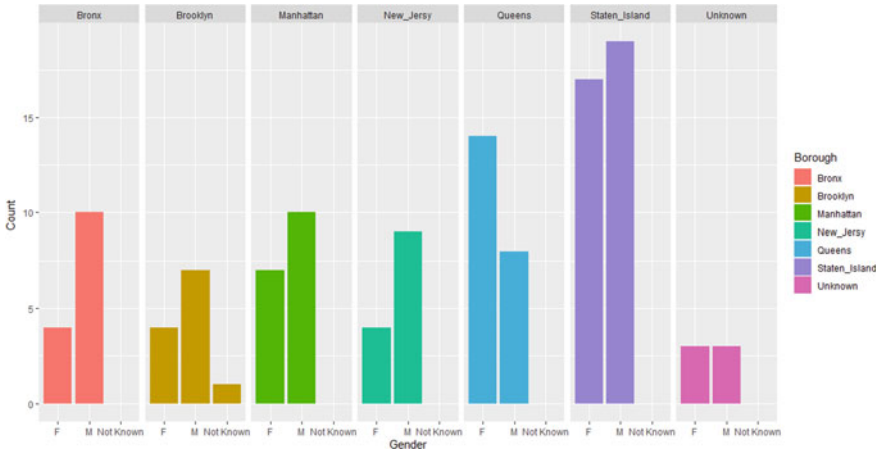


Fig. 6 Gender distribution with respect to boroughs after cleaning

Fig. 7 Favorite dog breeds after data cleaning

Therefore, pug is the most favorite dog in whole NYC region of The United States (Fig. 7).

5 Conclusion and Future Work

Data mining is the process of discovering useful information in a large data repository [11]. This single phase requires number of pre-requisite activities to be followed in a sequence. In our work, we have covered all data mining activity level with inclusion of new activities to improve the knowledge discovery procedure. We carefully analyze the deep insights for data collection and its preprocessing units and suggested an algorithm for effective mining of textual dataset. This algorithm can also be useful in enhancing the over all data quality of any analytics system. The results section demonstrates each step of the proposed algorithm and shows the fruitful impact on dataset quality and knowledge discovery. This has capability of extension, if

any new abnormality is found in future. More explorations can be done on cleaning requirements of textual datasets using fusion of machine learning algorithms. It would be nice if a single sequence of this data mining algorithm gives best performance for each type of dataset.

References

1. Grzymala-Busse, J. W., Grzymala-Busse, W. J.: Handling missing attribute value. In O. Maimon & L. Rokach (Ed.), *Data mining and knowledge discovery handbook* (pp. 573–589). Springer.
2. Logan, C., Parás, P., Robbins, M., & Zechmeister, E. J. (2020). Improving data quality in face-to-face survey research. *PS: Political Science & Politics, 53*(1), 46–50.
3. Cai, L., & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal, 14*.
4. García, S., Ramírez-Gallego, S., Luengo, J., et al. (2016). Big data preprocessing: Methods and prospects. *Big Data Analytics, 1,* 9.
5. García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining*. Intelligent Systems Reference Library.
6. García, S., Luengo, J., & Herrera, F. (2016). Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. *Science Direct Knowledge-Based Systems, 98,* 1–29.
7. Oni, S., Chen, Z., Hoban, S., & Jademi, O. (2019). A comparative study of data cleaning tools. *International Journal of Data Warehousing and Mining, 15*(4), 48–65.
8. Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2020). Factors influencing effective use of big data: A research framework. *Information & Management, 57*(1), 103146.
9. Alasadi, S. (2017). Review of data preprocessing techniques in data mining. *Journal of Engineering and Applied Sciences, 12,* 4102–4107.
10. Liu, Q., Xiao, F., Zhao, Z. (2020). *Grouting knowledge discovery based on data mining. Tunnelling and Underground Space Technology, 95.* ISSN 0886-7798.
11. Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining* (2nd ed.). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.
12. NYC Open Dataset. (2018). https://data.cityofnewyork.us/browse?q=NYC+Dogs+Licenses.
13. Postal Codes of New York City. (2019). https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoods.htm.
14. Postal Codes of The United States of America. (2019). https://www.postalpinzipcodes.com/Postcode-USA-United-States-ZIP-Code-8527-Postal-Code.
15. Cordón, I., Luengo, J., García, S., Herrera, F., & Charte, F. (2019). Smartdata: Data preprocessing to achieve smart data in R. *Neurocomputing, 360,* 1–13.

A Smart App for Pothole Detection Using Yolo Model

Rajshekar Hiremath, Komal Malshikare, Manish Mahajan, and Radhika V. Kulkarni

Abstract Pothole is the structural failure on the road, which causes accidents. In India, due to an increase in transportation, the number of mishaps because of potholes has additionally expanded. In this way, for diminishing the loss of human life because of potholes a few techniques has been conceived to identify the potholes utilizing sensors. These techniques are exorbitant and inefficient. So we have structured a savvy approach which utilizes cell phones with camera and GPS sensors. Here, we are using "YOLO object detection" algorithm to detect potholes. The application detects the location of a pothole. The users can upload images of potholes in their area. After uploading, the YOLO algorithm validates the given image. Then, the location of the pothole is displayed on the map. Civic authority in that area can repair the potholes. So in this strategy, we are executing the tech-savvy and sustainable answer for pothole identification, and this technique can successfully identify street road conditions utilizing the cell phone.

Keywords Convolution neural network · Global positioning system · Image classification · Object detection · Roads safety · Pothole detection · YOLO model

R. Hiremath (✉) · K. Malshikare · M. Mahajan · R. V. Kulkarni
Department of Information Technology, Pune Institute of Computer Technology, Pune, India
e-mail: rajshekar2198@gmail.com

K. Malshikare
e-mail: kg.malshikare@gmail.com

M. Mahajan
e-mail: manishsm11@gmail.com

R. V. Kulkarni
e-mail: radhikavikaskulkarni@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_16

1 Introduction

Transportation is the leading sector in every country. Due to the increasing number of vehicles, the likelihood of mishaps in India has expanded. Road surface conditions affect transport safety and driving comfort. The user should be aware of the road conditions for safety purposes. A few strategies [1–6] have been proposed, however, are inefficient and expensive. There is no computerized framework to recognize the potholes. Hence, civic authority and citizens both have to face the challenges [1]. The severity caused by the potholes is largely unnoticed. Daily, death of ten people is reported due to accidents caused by potholes in India [7]. In 2017, about 3597 deaths and 25,000 injured were reported in accidents due to the bad road conditions [8]. Over 9300 deaths have been reported due to potholes in the last three years [8]. This causes loss of human life and damage cost to the roads and the vehicles involved.

Hence, we have thought of a sustainable solution to counter this problem by developing a smart user-friendly app for people to detect the road surface conditions. Our solution is fast, accurate, inexpensive to operate, and utilizes a smartphone's camera and GPS. By applying the YOLO algorithm [9], the proposed method efficiently detects multiple potholes in the image and accordingly defines the road surface conditions.

2 Related Work

Many researchers have contemplated different techniques to distinguish the potholes. Some methods [1, 2] are very difficult to use, some [3, 4] have very low accuracy, and some [5] perform only in favorable conditions. Some methods [10, 11] give good accuracy but are inefficient due to high computation and expensive techniques.

The study in [1] proposes a low-cost sensor and CNN-based method for automatic pothole detection. This method is unable to detect the pothole image under the condition of illumination variation [1]. Some researchers describe sensor-based methodologies to detect the potholes [2, 5]. The study in [2] introduces road damage detection using ANN and accelerometer, with images captured through a smartphone mounted on the car with its GPS and accelerometer activated. This method detected and classified the three types of roadway anomalies and showed a high degree of accuracy, but did not detect potholes. Piao and Aihara [5] use sensors which present a vibration-based system for pothole detection. They are using mobile sensors that contain both accelerometer and GPS. Here, high error rate is noticed because this method is more physical than the digital one.

Some methods determined the road conditions according to the roughness of the road. The research work in [3] shows that the root mean square of the vertical component of the acceleration has a high correlation with IRI. Using this, approximate roughness of the road can be detected. Here, the limitation is that the parameter must be manually adjusted for each vehicle. Yagi [4] describes a spring and damper

model. It can automatically estimate vehicle parameters including the damping ratio and resonant frequency which can be used to detect the road roughness. This study estimates the road roughness index and identifies changing road conditions.

The research work in [11] suggests creating 3-D images from two 2-D images captured by two different cameras aligned in a specific manner. Hence, potholes can be detected based on their geometric shape. It provides high accuracy but it is highly dependent on the orientation of cameras on vehicles. Also, a lot of computational power is required to create initial 3-D images [11].

3 Proposed Methodology

3.1 Android Application and Google Maps API

In the proposed methodology, we are using an android application along with Google maps API. Android applications are widely used by people all over the world due to its easily accessible platform. This app is used in our method, as it is the most easily accessible device in one's pocket and can be used by a number of people widely over the world [2].

Google APIs permit the use of Google services. One such API is an embedded Google map. It can be accomplished by utilizing the Static maps API, Places API, or Google Earth API. This map API is utilized in our method. Google maps API helps us to integrate the maps into our system.

3.2 YOLO Model

YOLO stands for "you only look once." This is the fastest algorithm that even can be used for object detection in live stream video [9]. Due to its good accuracy and fast rate of image recognition, we deployed this algorithm to detect the potholes [9]. It is faster than other detection systems across a variety of detection datasets. It is famous because it achieves good accuracy while also being able to work in real-time and provides a frame rate of 45 fps [12]. It has a smooth trade-off between speed and accuracy [13]. The algorithm "only looks once" at the image in the sense that it requires only one forward propagation pass through the neural networks to make predictions [14].

3.3 Methodology

The proposed methodology consists of a smartphone application useful for citizens and civic authority to detect the road surface condition. It is a simple user-friendly

app that enables its user to tap pictures of potholes and upload them. This application is used to detect the potholes as well as view the locations of the potholes. Citizens register on the app using their email and username–password. Access permissions, like the location of the user (i.e., GPS), camera permissions are required. User logs into his account and taps the photos of the potholes in their area. When one submits the image, our application processes the request and sends it to our server. Within a few seconds, he is notified, the number of potholes detected in the image. If potholes have been detected, then it is stored in our database and displayed on the map for everyone to view. At the backend, to detect the potholes in the image, neural networks are being used. Specifically, we are using the YOLO algorithm to detect the potholes in the image. Civic authority registers on the app and he logs in using his credentials. Civic authority can view the potholes in his locality that have been reported by the users on the map. Accordingly, he repairs the potholes and updates the status of the work in our app. In this way, our app effectively works to detect potholes and easily repairs them.

For more effective classification, the potholes are categorized and given more priority according to their severity (or location), i.e., national highway, state highway, city rods, local roads, and number of potholes detected in one image or area. More the user base, more effective our app works. To increase our user base, we are providing credit points to citizens for each upload. Likewise, the civic authority will be rated and rewarded according to his work completed in a stipulated time. Our methodology consists of an android app, which is easy to use smart way to detect the potholes. This method has high speed and good accuracy in detecting potholes in the image. It also can efficiently detect multiple potholes in a single image.

3.4 Workflow of the App

1. Registration: Users can register on the app as a citizen or civic authority with their email and password.
2. Login: Users can log into the app by the credentials provided to him.
3. Access to permission: User then has to give access to permissions of GPS and camera.
4. Capture Image: User then has to open the camera, click a picture of the pothole, and submit it.
5. Processing and result: After submitting the image, within seconds the image is processed, and then, the user is notified the number of potholes detected.
6. Google Map: Users can access the map, integrated into our app, and view the markings/locations of the pothole in the area.
7. Logout.

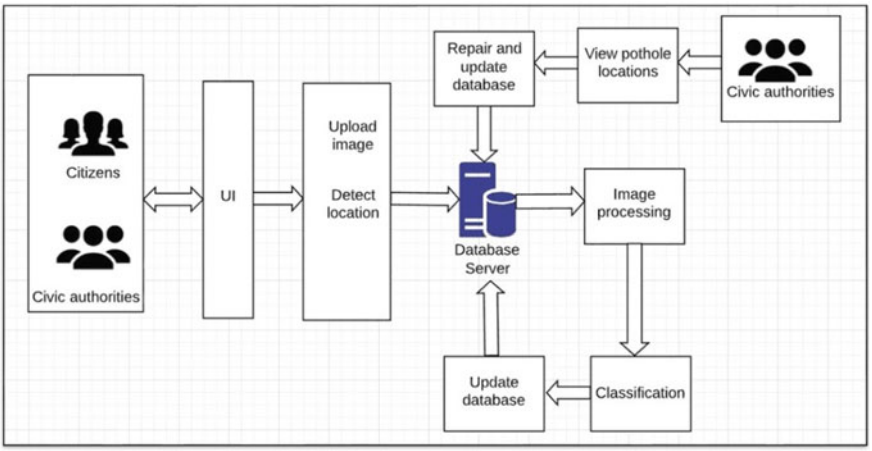


Fig. 1 System architecture

3.5 Architecture

The proposed system architecture consists of an android application, image processing model, and server to process in request and store data. We are trying to automate the system by providing real-time pothole locations (Fig. 1).

3.6 Pothole Detection Using YOLO Algorithm

Initially, we take a pre-processed pothole picture, and YOLO [9] is applied. The picture is isolated as frameworks of grids. We then isolate the picture into any number matrices, contingent upon the picture. Each grid undergoes classification and localization. The objectness or the certainty score of every matrix is found. In the event that there is no any pothole found in the framework, at that point the objectness and bounding box estimation of the network will be zero, or in the event that we have found a pothole in the lattice, at that point the objectness will be one, and the bounding box value will be its bounding value of the discovered object. To comprehend the YOLO calculation, it is important to build up what we need to foresee [14]. At last, we aim to predict a class of an object and the bounding box indicating object area. Each bounding box has four descriptors: Center of a bounding box (**bx**, **by**), width (**bw**), height (**bh**), value **pc** is corresponding to a class of an object (as: pothole or no pothole) [15].

Also, we need to anticipate the pc value, which is the likelihood that there is an object in the bounding box. As mentioned above, when working with the YOLO algorithm we are not searching for interesting regions in our image that could potentially contain an object. Rather, we are parting our picture into cells, utilizing a 19 ×

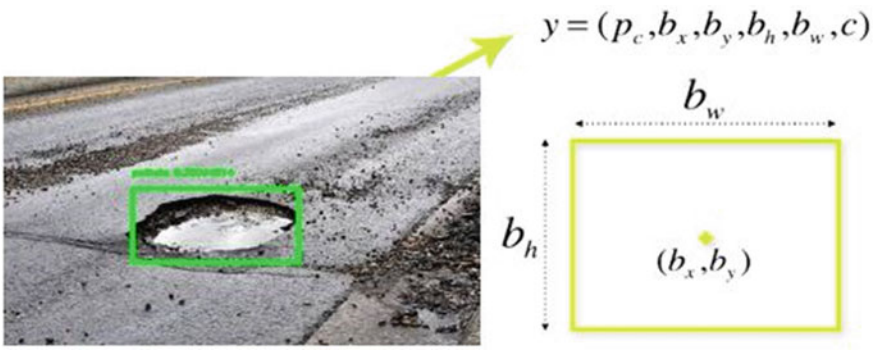


Fig. 2 Bounding box description

19 lattice. Every cell is liable for foreseeing five bounding boxes (in the event that there is more than one object in this cell). In this manner, we show up at an enormous number of 1805 bounding boxes for one picture. A large portion of these cells and bounding boxes won't contain an object. Most of these cells and bounding boxes will not contain an object. Therefore, we predict the value pc, which serves to remove boxes with low pothole probability and bounding boxes with the highest shared area in a process called non-max suppression [14] (Fig. 2).

4 Experimentation and Results

4.1 Pre-processing

Pothole images were collected from the kaggle pothole dataset [16]. These images are used for training and testing of our YOLO model. One thousand five hundred images were used with five-fold cross-validation. Images were pre-processed before training. Following are processing methods being used like checking annotations and labels of the data, checking bounding boxes, scale images to 80–120%, converting images to superpixel representation, horizontal and vertical flip if required, adding Gaussian blur to images, detect the edges, scale, and resize images if required.

4.2 Experimentation

For carrying this experiment, we have used pothole kaggle dataset [16]. About 1500 images are used for training our YOLO model. The model is trained with five-fold cross-validations. Table 1 shows the confusion matrix for the training and testing of 1500 images. By applying five-fold cross-validation, the accuracy of the proposed

Table 1 Confusion matrix of five-fold cross-validation

	Fold 1		Fold 2		Fold 3		Fold 4		Fold 5	
Class	Class 1 predicted	Class 2 predicted	Class 1 predicted	Class 2 predicted	Class 1 predicted	Class 2 predicted	Class 1 predicted	Class 2 predicted	Class 1 predicted	Class 2 predicted
Class 1 actual	160 (TP)	40 (FN)	164 (TP)	36 (FN)	165 (TP)	35 (FN)	172 (TP)	28 (FN)	176 (TP)	24 (FN)
Class 2 actual	40 (FP)	60 (TN)	35 (FP)	65 (TN)	33 (FP)	67 (TN)	26 (FP)	74 (TN)	19 (FP)	81 (TN)
Accuracy (%)	69.1		74.3		75.66		79.21		85.45	
Overall accuracy (%)					76.4					

Here,
Class 1: Pothole. *Class 2*: No Pothole
True Positive (*TP*): Observation is that there is a pothole and is predicted as a pothole
False Negative (*FN*): Observation is that there is a pothole but is predicted as not a pothole
True Negative (*TN*): Observation is that there is no pothole and is predicted as not a pothole
False Positive (*FP*): Observation is that there is no pothole, but is predicted as a pothole

Table 2 Number of images which detected the exact number of potholes

	Images detected with the accurate number of potholes	Images detected with inaccurate number of potholes than actual	Accuracy (%)
Fold 1	130	30	81.2
Fold 2	135	29	82.3
Fold 3	138	27	83.6
Fold 4	148	24	86.1
Fold 5	150	26	85.2
Total	701	136	83.6 (Average)

method comes out to be 76.4%. With an increase in training data, the model can improve its accuracy. Further, the proposed method also identifies the number of potholes in an image. It detects the exact number of potholes in an image with the accuracy of 83.6% (Table 2). Thus, along with good accuracy to detect potholes, it also shows better accuracy in detecting multiple potholes in an image. More number of potholes in an image describes the worse road condition.

The input image size for our system is set to 416 × 416 pixels with three channels of color as red, blue, and green. Our model is trained with object threshold of 0.5, and non-max suppression is set to 0.45. Maximum boxes per image are set to 18. After exceeding 18, the pothole is set on the high priority list. Confidence score is set above 80 to accurately detect the potholes in the image. The output of our model is presented with four parameters. (1) Location: Multidimensional array representing bounding boxes (2) Class: Indicating the class (Pothole or No pothole). (3) Score: Representing the probability that a class was detected. (4) Number of detections: Array of length 1 containing a value expressing the total number of potholes detected in an image. Refer (Fig. 3).

Various algorithms were studied and compared for object detection in the image. It is found that the YOLO algorithm has quicker handling time than any other algorithm.

Fig. 3 Multiple potholes detected in an image

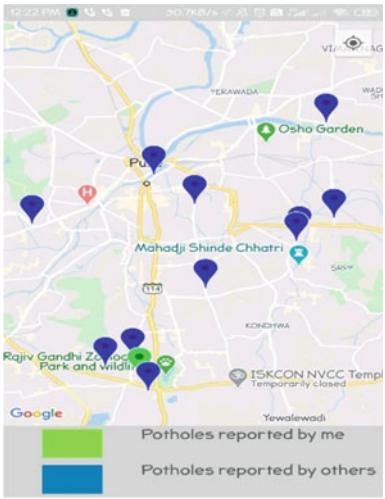


Fig. 4 Potholes location on the map

It could also be used for real-time object detection [13]. It is accurate in detecting the multiple potholes in the single images which is poor in the case of other algorithms. Our model gives an accuracy of 76% [14]. Hence, due to multiple pothole detection and low processing time, this is effective for real-time pothole detection. Utilizing results from gadgets running the application, we have delivered rich maps of the city with potholes and street conditions, as appeared in Fig. 4.

5 Conclusion and Future Scope

We studied different approaches for pothole detection. We have selected the best suitable YOLO algorithm for pothole detection. It detects the presence of potholes in an image with around 76% accuracy; in these correctly detected images, it counts the exact number of potholes with the accuracy of 83.6%. This app helps to detect fake images by making use of machine learning. Since the users directly feed the pothole images to the app, it reduces the burden of redundant storage and pre-processing. This mobile app also categorizes the road conditions depending on the presence of number of potholes. It helps the civic authorities to take remedial actions on priority basis.

This project can be further extended for other activities like comparing different road conditions and selecting the shortest route to the destination with the better road conditions. Further, we can add the feature to produce color-coded maps based on the severity of the road surface conditions. The efficiency of the model for recognizing potholes can be further enhanced by applying backward learning to make the system smarter with time. In future, this work can be extended to detect road conditions in all other cities.

References

1. Kumar, A., & Kumar, S. (2018). Road quality management using mobile sensing. In *International Conference on Innovations in information Embedded and Communication Systems.*
2. Mendis, A., Starzadnis, G., & Anoris, G. K. (2019). Real time pothole detection using android smartphones and accelerometers. In *2019 12th International Conference on ITS Telecommunications (ITST)* (pp. 668–672). IEEE.
3. Mohan, P., Padmanabhan, V. N., & Ramjee R. (2008). Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of 6th ACM Conference on Embedded Networked Sensor Systems* (pp. 323–336).
4. Yagi, K. (2016, September). A measuring method of road surface with road profiler. In *Proceedings of IEEE Transaction on Intelligent Transportation System* (Vol. 13).
5. Piao, B., & Aihara, K. (2019). Detecting the road surface condition by using mobile crowd sensing with drive recorder. In *IEEE 20th International Conference on Intelligent Transportation Systems.*
6. Kyriakou, C., Christodoulou, S. E., & Dimitriou, L. (2016). Roadway pavement anomaly classification utilizing smartphones and artificial intelligence. *Journal of the Transportation Research Board*.
7. https://www.theguardian.com/world/2018/jul/25/more-deadly-than-terrorism-potholes-responsible-for-killing-10-people-a-day-in-india.
8. https://www.indiatoday.in/india/story/over-9300-deaths-25000-injured-in-3years-due-to-potholes-1294147-2018-07-24.
9. Redmon, J., & Farhadi, A. (2020). YOLO9000: Better, faster, stronger. In *IEEE Conference on Computer Vision and Pattern Recognition.*
10. Lee, S. W., & An, K. E. (2018). Image processing based pothole detection system for driving environment. In *IEEE International Conference on Consumer Electronics (ICCE).*
11. Zhang, Z., Ai, X., & Chan, C. K. (2018). An efficient algorithm for pothole detection using stereo vision. In *IEEE International Conference on Acoustic, Speech and Signal Processing*.
12. Comparison: https://medium.com/@jonathan_hui/object-detection-speed-andaccuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359.
13. Divvala, S., Girshick, R., & Farhadi, A. (2017). You only look once: Unified, real-time object detection. In *Computer Vision-ECCV 2014 Workshops* (pp. 101–116). Springer.
14. Geethapriya, S., Duraimurugan, N., & Chokkalingam, S. P. (2019, February). Real-time object detection with Yolo. *International Journal of Engineering and Advanced Technology (IJEAT), 8*(3S). ISSN: 2249 – 8958.
15. YOLO: Real-Time Object Detection. https://pjreddie.com/darknet/yolov2/.
16. Kumar, A. Dataset for potholes. https://www.Kaggle.com/atulyakumar98/pothole-detection-dataset.

Medicare with Machine Learning and Deep Learning

Jinam Pankajbhai Mehta, Sanket Patel, Mukesh Chavada, Rashes Dobariya, and Vivaksha Jariwala

Abstract India is facing a scarcity of medical expertise, modern medical equipment, and hospitals. People in rural areas don't even have access to healthcare facilities. At the same time, country is struggling to fight with deadly diseases like malaria and pneumonia that causes thousands of deaths every year. Hence, it becomes the necessity to come up with an alternate solution that can eliminate the challenges posed by lack of enough medical professionals and health infrastructure. Artificial intelligence concepts machine learning and deep learning can be applied to automate the process of disease identification and diagnosis. The paper proposes the system named "Medicare" that operates in two parts. The first module allows the user to predict the potential diseases by providing the likely symptoms. The module works on the concept of machine learning. We feed the dataset comprises of various diseases and associated symptoms pairs. Both doctors and normal people can make use of this disease predictor. Doctors can use this to verify their diagnosis results with the outcome of the system. Normal people can use this to diagnose the disease at home. The system depicts the potential diseases with their probabilities and advises them whether to consult the doctors or not. The second module allows us to test whether a medical image is positive or negative. This module works on the concept of a convolutional neural network that performs the image processing on the image to

J. P. Mehta (✉) · S. Patel · M. Chavada · R. Dobariya · V. Jariwala
Department of Information Technology, Sarvajanik College of Engineering & Technology, Surat, Gujarat 395002, India
e-mail: jinammaheta0@gmail.com

S. Patel
e-mail: Sbpatel8484@gmail.com

M. Chavada
e-mail: mukeshchavada16@gmail.com

R. Dobariya
e-mail: rashesdobariya70@gmail.com

V. Jariwala
e-mail: vivaksha.jariwala@scet.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_17

test the image. Little bit modification will allow the person to test several images at the same time, hence will save time and cost.

Keywords Machine learning · CNN · Deep learning · Medical

1 Introduction

Though India is the second-most populous country, the number of doctors available is not sufficient. The government of India's data says that India has less than one doctor per thousand people that is even lower than the world health organization (WHO) standard [1].

Another problem is medical checkup is a very time-consuming process and expensive as well. They have to make an appointment first and consult the doctor even though there is no serious issue. It has been seen that sometimes doctors have conflicting opinions when it comes to the diagnosis of a patient. These wrong opinions of doctors may lead to misdiagnosis. That is, where our system comes into the picture. It provides reliable predictions of potential diseases. It uses the supervised machine learning classification algorithm. The classification algorithm finds the probabilities for each class for a given vector of features and classifies the given input into the class with the highest probability [2, 3].

At the same time, that medical expertise is not available in rural India leads to unequal access to public health in India [4]. If we look at some diseases that are causing a large number of deaths, one of them is malaria, especially in South Asia. As per guideline from WHO protocol, to detect malaria, it requires examination of blood smear at 100 × magnification, where people count manually the number of cells that contain parasite out of 5000 cells [5]. Another deadly disease is pneumonia that also requires the radiologist to carefully examine x-ray images. As we all know skilled people are scarce, it is unlikely that we will able to diagnose all cases. Here comes our second module that performs image processing and makes use of deep learning to detect the disease from images faster and more accurately. Deep learning is a subset of machine learning that is specialized in image classification. Algorithms automatically identify the important features from training samples eliminating the requirement of manually feeding features. There are several types of neural networks, but convolutional neural network (CNN) is efficient for classifying images [6].

2 Literature Review

The author [7] emphasizes the success of predictive tasks in a wide range of areas that can be replicated in the medical field as well. He demonstrates the use of the machine learning task by creating three predictive models for cancer diagnosis. These algorithms are regularized general linear model regression (GLMs), support vector

machines (SVMs) with a radial basis function kernel, and single-layer artificial neural networks.

The author [8] reviews the computer-assisted analysis of images in the area of medical imaging. He says that recent advancements in deep learning help to identify the pattern and classify the medical images.

The author [9] discusses the problem faced by the health care sector and how deep learning can help to solve the problem. He mentions the complexity in analyzing medical image manually and how deep learning can dominate the future health-related apps.

3 Proposed Approach

3.1 Plan of Work

- Collect datasets: This step involves importing raw data from various online dataset repositories.
- Import libraries: This step involves the installation of various required python libraries like Keras, pandas, NumPy, and many more.
- Data preprocessing: This step involves converting data (excel datasheet or image dataset) into a format appropriate for processing by algorithms.
- Model training: This step involves the training of the model on the training dataset and testing on the validation dataset. We will iterate over this step until we get the desired accuracy.
- Model Testing: Here, we will evaluate our model on unseen data and will check if we are getting the desired performance. If not, then we will make appropriate changes (Fig. 1).

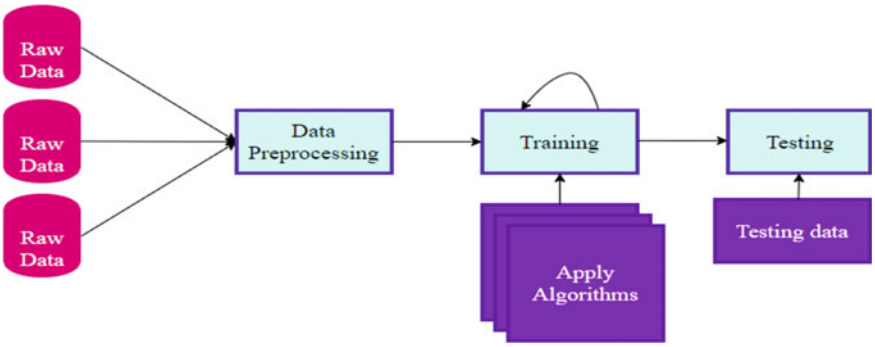


Fig. 1 Plan of work for proposed system. It shows steps to be performed for creating the model that will diagnose the disease

3.2 Disease Prediction Using Machine Learning

Disease predicator can make use of various machine learning algorithms to predict the disease employing probability and statistics. One of them is naïve Bayes that calculates the probability of each disease and returns one with the highest probability. Naïve Bayes assigns the most likely class to the given feature vector. The naïve Bayes classifier simplifies the learning by assuming that features are independent [10]. The probability that given input feature vector X belongs to class C is,

$$P(X|C) = \pi_{i=1}^{n} P(X_i|C)$$

To improve prediction accuracy, we have used a technique named association rule mining that uses the Apriori algorithm. Association rule mining also known as market basket analysis generates frequent itemset from the dataset [11, 12]. Each row in the dataset represents disease and its corresponding symptoms. The number of frequent items generated depends upon the term minimum support. After generating frequent itemset, association rules are extracted that satisfy the minimum confidence and lift requirement. These rules are useful in suggesting the frequently occurring symptoms based on the user's inputs.

3.3 Malaria and Pneumonia Identification Using CNN

Convolutional neural network is a deep learning algorithm used for the classification of the images. It assigns the weights and biases to the network during the training of the model. Each input image during the training process passes through a series of convolutional layers and filters, polling layers, fully connected layers, and in the output layer, either softmax or sigmoid function is applied to classify the image [13, 14].

Convolution layer is the first layer in the network that performs feature extraction from the image. It takes two argument image matrix and filter. Convolution of an image with different filters allows us to detect import features at each layer like edge detection, blur the image, grayscale the image. The polling layer allows us to reduce the number of parameters while retaining important information. Most popular is max polling that extracts the maximum from a given feature map.

The model that we have built comprised of 11 layers with 5 layers that are part of the fully connected network, and rest is convolutional layers. All layers use ReLU activation function except the output layer. The output layer uses a sigmoid function to classify the input in either a positive or negative category. The model is trained using Adam optimizer with loss function binary cross-entropy. The model is trained with a batch size of 32 and 5 epochs. We have also performed image augmentation to improve accuracy as it covers a wide range of use cases. We have used the EarlyStopping technique to halt the training of the model when validation accuracy starts decreasing.

4 Result and Analysis

4.1 Technology Stack

We have used PyCharm and Visual Studio for code development. Python, JavaScript, and HTML are mainly used with the flask framework to create the python web app and bootstrap for a user-friendly interface. To simplify the development process and to create a robust app, various libraries such as jQuery, Keras, Sklearn, Pandas, NumPy, Matplotlib, TensorFlow have been used.

4.2 Association Rules Generation

Total number of association rules generated depends upon the parameters such as minimum support, minimum confidence, and lift. Total number of rules changes as these parameters change (Figs. 2 and 3). These three are very important for generating relevant and useful association rules. The support term says that how frequent itemset (symptoms) in the database is. It is measured by the proportion of the transactions or rows in which itemset appears. Confidence term says that how likely itemset Y will appear if itemset X is there. It is measured by the proportion of transactions with itemset X, in which itemset Y also appears. Lift says that how likely item Y appears when item X appears while considering the support for item Y that means how popular item Y is. The lift parameter increases the trustworthiness of the association rule.

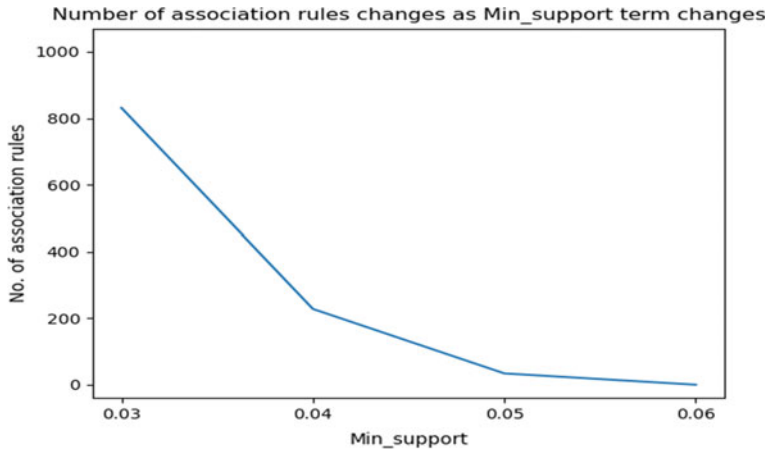


Fig. 2 Minimum support impact: It shows how number of rules generated by algorithms decreases as minimum support term increases while keeping other terms constant

Fig. 3 Minimum confidence impact: It shows how number of rules generated by algorithms decreases as minimum confidence term increases while keeping other terms constant

4.3 CNN Model for Malaria Identification

CNN model that we have built achieved 96.33% accuracy on training data and 96.39% accuracy on validation data after 5 epochs (Fig. 4). The classification report points out an f1-score of 0.99 for the infected label and 0.98 for the uninfected label (Table 1).

Training accuracy, validation accuracy, training loss, and validation loss for each epoch is shown in Table 2 that compares the performance of our CNN model with the MobileNet CNN model.

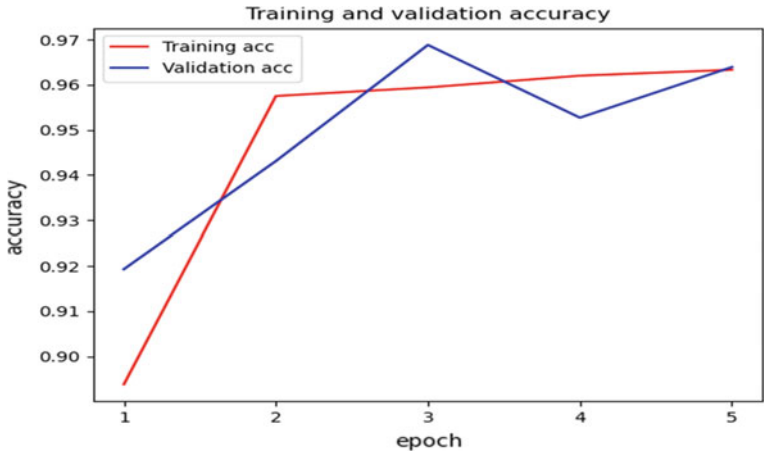


Fig. 4 Training and validation accuracy: It shows how model performs on training data and validation data after each epoch

Table 1 Classification report

	Precision	Recall	f1-score	Support
Infected	0.99	0.98	0.99	173
Uninfected	0.97	0.99	0.98	141
Accuracy	–	–	0.98	314
Macro average	0.98	0.98	0.98	314
Weighted average	0.98	0.98	0.98	314

Table 2 Accuracy and loss comparisons between our model and MobileNet model

Epochs	Self-made model				MobileNet model			
	Loss	Accuracy	Validation loss	Validation accuracy	Loss	Accuracy	Validation loss	Validation accuracy
1	0.2455	0.8938	0.2099	0.9193	0.3872	0.8178	0.2871	0.9233
2	0.1312	0.9575	0.1722	0.9431	0.1711	0.9442	0.1486	0.9615
3	0.1222	0.9594	0.1185	0.9688	0.1520	0.9531	0.1436	0.9567
4	0.1122	0.9620	0.1355	0.9527	0.1385	0.9562	0.1596	0.9527
5	0.1077	0.9633	0.1261	0.9639	–	–	–	–

5 Conclusion and Future Work

5.1 Conclusion

We have trained the machine learning model to predict the disease from symptoms and used association rule mining for more accurate results. Apart from that, we have used CNN for the identification of deadly disease malaria and pneumonia, while traditional identification approaches require lots of medical expertise, finances, and time. We have also trained the MobileNet model to compare accuracy and loss with our model. We hope the web app that we have developed will drastically reduce the cost and time to diagnose the patient and will also eliminate the requirement of medical expertise to analyze the medical images. It will be more helpful in remote areas, where medical resources are scarce. But there are some limitations to the system. One is the user may not be able to understand some of the medical terms results in insufficient details. Another major drawback is that it can classify only one image at a time.

5.2 Scope of Future Work

The system can be extended to predict the variety of diseases by training the model on a huge dataset that includes several features. Some of the features like body temperature, blood pressure can be measured directly by medical devices. This will make the process partially automated by taking readings from medical devices and input it to the predictor.

References

1. Indian Express article. https://indianexpress.com/article/india/less-than-one-doctor-for-1000-population-in-india-government-tells-lok-sabha-4760892/.
2. Santafe, G., Inza, I., & Lozano, J. A. (2015). Dealing with the evaluation of supervised classification algorithms. *Artificial Intelligence Review, 44*(4), 467–508.
3. Dutta, N., Subramaniam, U., & Padmanaban, S. (2020, January). Mathematical models of classification algorithm of Machine learning. In *International Meeting on Advanced Technologies in Energy and Electrical Engineering* (Vol. 2019, No. 1, p. 3). Hamad bin Khalifa University Press (HBKU Press).
4. ncbi report. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4621381/.
5. WHO report. https://www.who.int/malaria/publications/atoz/who-severe-malaria-tmih-supplement-2014.pdf.
6. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M. … Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42,* 60–88.
7. Sidey-Gibbons, J. A., & Sidey-Gibbons, C. J. (2019). Machine learning in medicine: A practical introduction. *BMC Medical Research Methodology, 19*(1), 64.
8. Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering, 19,* 221–248.
9. Razzak, M. I., Naz, S., & Zaib, A. (2018). Deep learning for medical image processing: Overview, challenges and the future. In *Classification in BioApps* (pp. 323–350). Cham: Springer.
10. Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In: *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence* (Vol. 3, No. 22, pp. 41–46).
11. Zeng, M., Xiong, Q., & Li, K. (2018, May). Design and implementation of an improved apriori data mining algorithm. In *8th International Conference on Social Network, Communication and Education (SNCE 2018).* Atlantis Press.
12. Gopakumar, G. P., Swetha, M., Sai Siva, G., & Sai Subrahmanyam, G. R. K. (2018). Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner. *Journal of Biophotonics, 11*(3), e201700003.
13. Liang, Z., Powell, A., Ersoy, I., Poostchi, M., Silamut, K., Palaniappan, K. … Huang, J. X. (2016, December). CNN-based image analysis for malaria diagnosis. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 493–496). IEEE.
14. Ye, Y., & Chiang, C. C. (2006, August). A parallel apriori algorithm for frequent itemsets mining. In *Fourth International Conference on Software Engineering Research, Management and Applications (SERA'06)* (pp. 87–94). IEEE.

A Comparative Study of Wireless Technologies Coexistence Mechanisms in IoT: A Survey

Marius Dasylva Mendy and Youssou Faye

Abstract The new communication model, Internet of Things (IoT) leads in the deployment of multiple and heterogeneous technologies in same frequency bands: the unlicensed bands. With the increase of connected devices, this cohabitation generates a lot of interferences which are the main source of performances degradation in wireless networks. In literature, coexistence of heterogeneous networks has been studied in many situations, especially in the WPAN (Wireless Personal Area Network) and WLAN (Wireless Local Area Network) colocation on 2.4 GHz band and in the spectrum sharing between Wi-fi and LTE (Long Term Evolution) on 5 GHz bands and not enough in IoT framework. The goal of this paper is to review wireless networks coexistence techniques developed on the MAC layer, then make a coexistence methods classification, bring a comparative analysis of these methods performances and finally study their adaptability in IoT context, regarding the applications, topologies, density and traffic.

Keywords Internet of things (IoT) · MAC layer · Wireless coexistence methods

1 Introduction

The need to control and interact with the environment leads to a new communication paradigm: the Internet of Things. This phenomenon induces many and various devices to be connected for services optimization and to facilitate daily life. Connected devices in the world are estimated to be billions in the next years [1]. This density and heterogeneity make various communication technologies to share the unlicensed bands ISM or U-NII, in the same or near area [2–7]. This coexistence makes interference which can strongly affect the operation of each technology. So the

M. D. Mendy (✉) · Y. Faye
University Assane Seck of Ziguinchor, Ziguinchor, Senegal
e-mail: mdasylva@univ-zig.sn

Y. Faye
e-mail: yfaye@univ-zig.sn

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_18

problem of coexistence in heterogeneous systems takes place in unlicensed bands. Although some researches focused on spectrum sharing, they are generally oriented to particular scenarios. In the context of the IoT, this problematic must cope with its specific characteristics. Indeed, this concept uses various wireless communication technologies on short (WPAN) or long range (LPWAN) to perform different applications types which run on different topologies and bit rates. Coexistence techniques specified in communication technologies are based on particular scenarios. Other techniques improve them and often lie on physical layer where techniques are rather manual, need additional hardware requirements and are often restricted [1]. In fact, since transmission comes from MAC layer, efficient coexistence technique should be mainly supported by this layer and as need be completed by physical process to deal with residual interferences. The purpose of this article is threefold: it reviews MAC based coexistences techniques for wireless communications technologies in the unlicensed bands and classify them in methods oriented in different approaches, analyzes performances of these methods in terms of various criteria and study their IoT context adaptability to help their choice according to the IoT context: application, topology, density and traffic.

The rest of this paper is organized as follow: Sect. 2 presents the related works and reviews the coexistence techniques; Sect. 3 describes applications in the IoT context. In Sect. 4 we make classification of coexistence methods based on different approaches. Then in Sect. 5, we make comparison of coexistence methods considering some performance criteria. To give orientations, Sect. 6 presents an analysis based on methods adaptability for IoT applications. Finally Sect. 7 concludes the paper and gives some perspectives.

2 Related Works and Review of Coexistence Techniques

Coexistence in unlicensed bands is studied in many works and generally lies on particular scenarios especially on Wi-fi and bluetooth colocation [3, 8] or in a Wi-fi and zigbee area [9, 10] on 2.4 GHz band. Recently, with the unlicensed LTE, many researches on its coexistence with Wi-fi which already occupies 5G bands are done. Surveys like [11, 12] are provided for this purpose. Kalaa and all in [13], review topics that encompass wireless coexistence with emphasize on medical devices. In the IoT specific domain, especially in LPWAN like LoRa and Sigfox, coexistence techniques are generally based on limitation of channel usage or communications on narrow/ultra narrow bands. Residual interferences are often resolved in physical layer like [14]. Fadeyi and all review in [1] possible optimization techniques on LPWAN and compare their merits and demerits. However these techniques which are focused on physical layer are often static, not scalable and costly. An efficient coexistence must be supported since MAC layer to better tackle harmful interferences. Works argue that coexistence techniques are based on physical, temporal or frequency separation [13, 15, 16]. Physical techniques involve Transmit Power Control (TPC) mechanism, used in IEEE802.11h standard [17] to manage Wi-fi coexistence with

radar system on 5 GHz band or in Wi-fi wide deployment [18]. Frequency techniques can be dynamic like Dynamic Frequency Selection (DFS) [17] in a cognitive radio approach and AFH (Adaptive Frequency Hopping) [3]. Its static form is FHSS used as Bluetooth access method or traditional FDMA-based techniques like UNB (Ultra Narrow Band) used in LPWAN [2]. Temporal techniques involve TDMA-based techniques like AWMA (Alternating Wireless Medium Access) in [3] specified for 802.11b and bluetooth colocation and bitmap scheduling scheme. Duty cycling used in LTE system for coexistence with Wi-fi on 5 GHz band [12] and in LoRa and Sigfox networks, LBT (Listen Before Talk) and CSMA/CA respectively used in LTE-LAA [19] and Wi-fi technologies to coexist with each other are non synchronized temporal techniques.

As a prelude to the study of methods adaptability to the IoT context, we present in the next section IoT context.

3 IoT Context

The main purpose of IoT is resources optimization and daily life simplification. For that, several applications are concerned but one can group them into three main categories [20]: (1) real time which is delay sensitive like health applications and topology is generally in star to urgently relay collected data to a user entity located out of the network, (2) device interaction which focuses on direct communications between connected devices like industrial applications, in a peer-to-peer topology and (3) data analysis, intended for storage and analysis in a star topology. These applications are rather reliability demanding and involve domains like smart city. In this article, we respectively name these categories, real time, interactive and periodic applications. Beyond this categorization, IoT applications are originally low rate but new ones emerge which need high rates, like audio and video applications.

4 A Classification of Coexistence Methods

Coexistence techniques reviewed above are often insufficient to provide efficient coexistence. So they are rather bricks of coexistence solutions that we can divide into two main approaches: self control or cooperative. Methods classification and their main characteristics are presented in Table 1.

4.1 Self Control Approach

In this approach, each network take into account that spectrum is shared with other network technologies and occupies it accordingly. However no collaboration is

Table 1 Comparison of coexistence methods

Coexistence methods						
Approach	Self control				Cooperative	
Methods	DC	NBC	CL	CR	DistC	CC
Coexistence	Intra/inter	Intra/inter	Intra/inter	Inter	Intra/inter	Intra/inter
Remaining interference	Yes	Yes	Yes	Yes	Yes(fewer)	no
Channels	Mono	Multi	Mono	Multi	Mono/multi	Mono/multi
Benefits	Energy consumpt.	Simultaneous transmissions	Implementation	Environ. awareness	Environment Knowledges	
Limits	High traffic	High bitrates	Asymetric detection	Channel choice	Overload, delay	Overload, delay, central unit
Key technologies	Lora, Sigfox, LTE-U	Sigfox, NB-IoT	LTE-LAA, Wi-fi, Zigbee	Wi-fi	–	–

performed. Four methods can be discerned. The first, Duty cycling (DC) [12, 21] which restricts transmissions possibility. It can deal with performance optimization by adapting to the surrounding context. The second, Channel-listening (CL) scan channel before transmissions. It involves LBT [9, 12], and CSMA/CA [10, 22], mechanisms which can naturally manage heterogeneous coexistence. However respective protocols must modify some features, [9] to ensure effective coexistence. The third, Narrow-band Communications (NBC) use small bandwidths to cope with interference and allow simultaneous transmissions [2], however sub-bands overlapping is frequent but frequency hopping principle can be used to avoid interference. The fourth, Cognitive radio (CR) opportunistically uses channels. It involves changing channel or reducing transmit power if transmission is sensed [23, 24]. To tackle channel choice issue, IA techniques like game theory models [25, 26] are often used to model channel choice as a non cooperative repeated game.

4.2 Cooperative Approach

To overcome lack of environment knowledge induced in self control methods, the cooperative approach introduces information exchange between networks so that radio resource sharing is more effective and fair. Cooperation can be distributed (DistC) or centralized (CC). In the former case, direct interactions between networks technologies are required based on message passing [27], on auction negotiation [28], on physical free medium signal [3], on energy level pattern for neighboring identification [12] or in cooperative game framework [29]. In the last case, central entity coordinates exchanges and performs resources sharing, in local or in cloud, to facilitate effective information sharing, often via a database [30].

5 Methods Analysis and Comparison

Table 1 presents a global comparison of coexistence methods whereas Fig. 1 illustrates methods performance next to some criteria. Self control methods because of their autonomous behavior are less scalable and fair and overlook some interference which reduce throughput. CR and NBC improve it thanks to multicanal and transmit power reduction for CR. Cooperative methods manage interference more effectively and fairly thanks to exchanged networks information and thus maximize throughput. Therefore they generate energy consumption, communication delay and bandwidth waste. They better support systems density. DistC is more flexible because the process is not centralized. However, it converges slower. Figure 2 shows that cooperative methods can better manage high density and high traffic scenarios. But it is hard in some cases for heterogeneous systems to interact or to have a central entity.

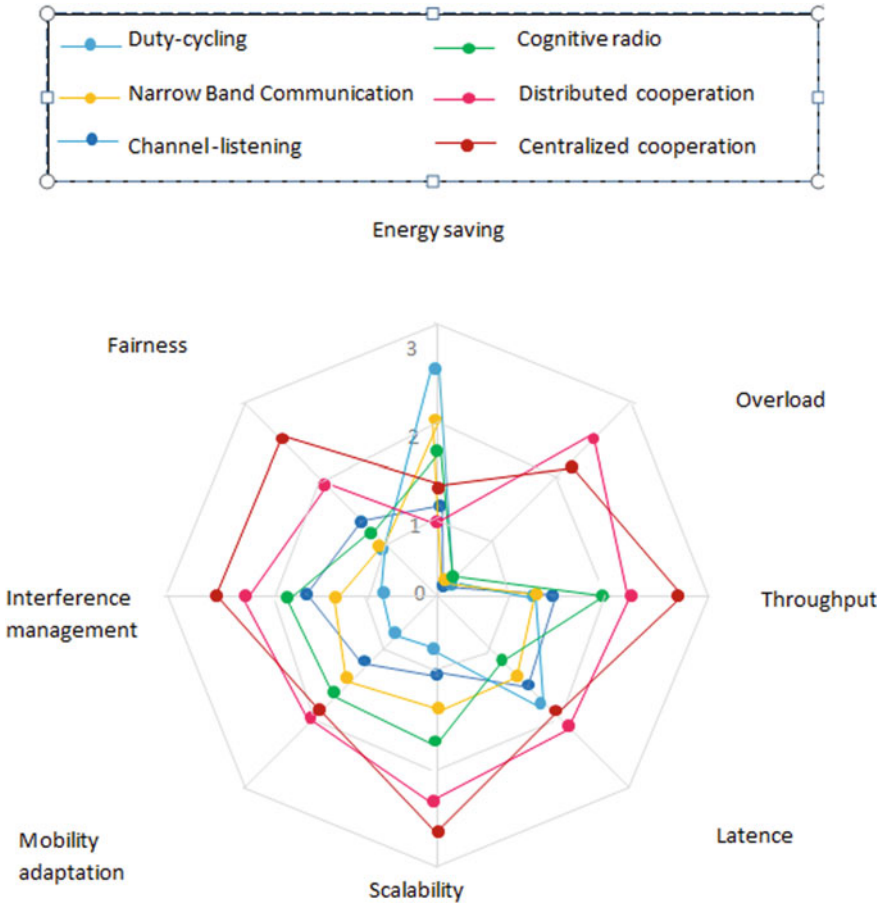


Fig. 1 Comparison of methods performances

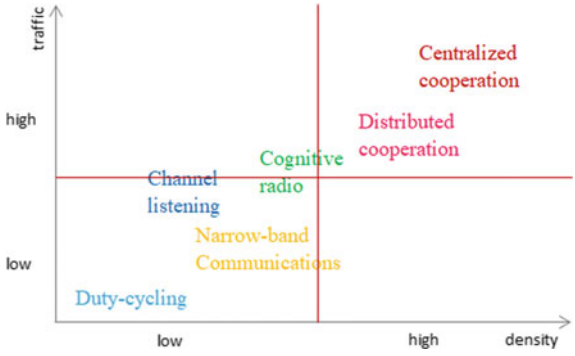


Fig. 2 Methods adaptation to traffic and density context

Table 2 Methods choice next to application type and context

Application type	Real time (health, supply chain)			Interactive (Industry, Home automation)		Periodic (Smart city, Smart automation)		
Topology	Star			Peer-to-peer		Star		
Density	Low		High	Low	High	Low		High
Traffic	Low	High	Low/high	Low/high	Low/high	Low	High	Low/high
Adapted methods	NB CL	CL	CR	CL	CR, DistC, CC	DC	DistC, CC	

6 Coexistence Methods Suitability for IoT Context

In this analysis, we consider an isolated IoT environment like health environment or a home automation system, which perform one kind of application. Let's remember that we distinguish three typical IoT applications: real time, interactive and periodic. Thus we address the adaptability issue from this observation: each kind of application generally runs in a typical communication topology: real time and periodic applications often run in star communication topology whereas interactive applications run in the most case in peer-to-peer topology. Now, communication topology, range, traffic type and density represent a context. So different IoT contexts are possible and previously cited methods are differently adapted to them. We elect the adapted methods according to this process: from the application type and its typical communication topology, we deduce the methods which cope with the application type requirements, and select which are applicable on the topology of this application, considering density and traffic parameters. We represent this process in a layered architecture illustrated in Fig. 3 and we show the methods and contexts matching in Table 2.

With its simultaneous communications which reduce waiting times, NBC can cope with real time applications, in low traffic environments. As for CL, its adaptation to these applications, in low density and traffic is due to its random access principle. For dense and high traffic contexts and despite the less scalability of signal mitigation

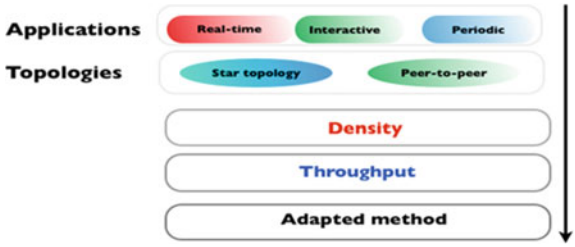


Fig. 3 Methods adaptation to traffic and density context

and overlapping channels, CR can, with adapted techniques, face up to real time applications. Thanks to its cyclic activity, DC is well suited to periodic applications in low density and traffic whereas high traffic and density are well managed by DistC and CC.

7 Conclusion

In IoT, connected devices and their various communication technologies need for adapted coexistence methods. In this paper we have reviewed coexistence techniques, classify the methods in distinct approaches and study their suitability in IoT contexts. Thus we saw that self control approach, by its independent nature, is well suitable for systems with no interaction. However due to the lack of environment knowledge, it overlooks some interference or reduces performance trying to avoid them. Cooperative approach is more efficient and fair but needs a central equipment or systems interactions which is often difficult in typical IoT environment. However when topology is a unique star, interactions are possible because various technology modules are located in the same equipment. Moreover we found that methods are differently adapted to the IoT context. Therefore, for each kind of IoT application and coexistence topology, we respectively deduce the most suitable methods. Further it will be interesting to consider an environment where multiple technologies run various applications with different requirements and forming a complex coexistence topology. Energy consumption or technologies status (aggressiveness) can also be considered to meet devices and networks states.

References

1. Fadeyi, J., Markus, E. D., Abu-Mahfouz, A. M. (2019). Technology coexistence in LPWANs-A comparative analysis for spectrum optimization. In *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, Vancouver (pp. 2244–2249).
2. Hattab, G., & Cabric, D. (2018, October). Spectrum sharing protocols based on ultra-narrowband communications for unlicensed massive IoT. In *2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Seoul (pp. 1–10).

3. ieee-recommended-practice-for-information-technology-local-and-m.pdf.
4. Masoudi, M., Azari, A., Yavuz, E. A., Cavdar, C. (2018, mai). Grant-free radio access IoT networks: Scalability analysis in coexistence scenarios. In *2018 IEEE International Conference on Communications (ICC)*, Kansas City, MO (pp. 1–7).
5. Sagari, S., Seskar, I., & Raychaudhuri, D. (2015). Modeling the coexistence of LTE and WiFi heterogeneous networks in dense deployment scenarios. In *2015 IEEE International Conference on Communication Workshop (ICCW)*, London (pp. 2301–2306).
6. Ancans, A., et al. (2019). Bluetooth low energy throughput in densely deployed radio environment. In *2019 23rd International Conference Electronics*, Palanga (pp. 1–5).
7. La, Q. D., Nguyen-Nam, D., Ngo, M. V., Hoang, H. T., & Quek, T. Q. S. (2018, December). Dense deployment of BLE-based body area networks: A coexistence study. *IEEE Transactions on Green Communications and Networking, 2*(4), 972–981.
8. Mourad, A., Muhammad, S., Al Kalaa, M. O., Hoeher, P. A., & Refai, H. (2017). Bluetooth and IEEE 802.11n system coexistence in the automotive domain. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, San Francisco (pp. 1–6).
9. Kim, C. K., Yang, C. S., & Kang, C. G. (2016). Adaptive Listen-Before-Talk (LBT) scheme for LTE and Wi-Fi systems coexisting in unlicensed band. In *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas (pp. 589–594).
10. Kim, Y., Lee, S., & Lee, S. (2015) Coexistence of ZigBee-based WBAN and WiFi for health telemonitoring systems. *IEEE Journal of Biomedical and Health Informatics, 20*(1), 222–230.
11. Ben Hafaiedh, H., El Korbi, I., Saidane, L. A., & Kobbane, A. (2017). LTE-U and WiFi coexistence in the 5 GHz unlicensed spectrum: A survey. In *2017 International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN)*, Paris.
12. Maglogiannis, V., Naudts, D., Shahid, A., Giannoulis, S., Laermans, E., & Moerman, I. (2017, août). Cooperation techniques between LTE in unlicensed spectrum and Wi-Fi towards fair spectral efficiency. *Sensors, 17*(9), 1994.
13. Kalaa, M. O. A., Guag, J., & Seidman, S. J. (2018). An outlook on wireless coexistence with focus on medical devices. *IEEE Electromagnetic Compatibility Magazine, 7*(3), 60–64. https://doi.org/10.1109/memc.2018.8479340.
14. Reynders, B., Meert, W., & Pollin, S. (2016, mai). Range and coexistence analysis of long range unlicensed communication. In *2016 23rd International Conference on Telecommunications (ICT)*, Thessaloniki, Greece (pp. 1–6).
15. Chew, K.-M., & Kelly, C. (2018, juin). RF co-existence testing for IoT wireless medical devices. In *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Rome (pp. 1–6).
16. Seidman, S., & LaSorte, N. (2014). An experimental method for evaluating wireless coexistence of a Bluetooth medical device. *IEEE Electromagnetic Compatibility Magazine, 3*(3), 49–54. https://doi.org/10.1109/memc.2014.6924328.
17. ieee-standard-for-information-technology–telecommunications-and.pdf.
18. Gandarillas, C., Martin-Engenos, C., Pombo, H. L., & Marques, A. G. (2014). Dynamic transmit-power control for WiFi access points based on wireless link occupancy. In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, Istanbul (pp. 1093–1098).
19. LTE Release 13.pdf.
20. Díaz, M., Martín, C., & Rubio, B. (2016, mai). State-of-the-art, challenges, and open issues in the integration of Internet of things and cloud computing. *Journal of Network and Computer Applications, 67,* 99–117. https://doi.org/10.1016/j.jnca.2016.01.010.
21. Ferré, G., & Simon, E. (2018). An introduction to Sigfox and LoRa PHY and MAC layers, p. 7.
22. Ishida, S., Tagashira, S., & Fukuda, A. (2015, déc.). AP-assisted CTS-blocking for WiFi-ZigBee coexistence. In *2015 Third International Symposium on Computing and Networking (CANDAR)*, Sapporo, Hokkaido, Japan (pp. 110–114).
23. Guan, Z., & Melodia, T. (2016). CU-LTE: Spectrally-efficient and fair coexistence between LTE and Wi-Fi in unlicensed bands. In *IEEE INFOCOM 2016—The 35th Annual IEEE International Conference on Computer Communications*, San Francisco (pp. 1–9).

24. Khan, I., & Harri, J. (2017, juin). Can IEEE 802.11p and Wi-Fi coexist in the 5.9 GHz ITS band ? In *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Macau, China (pp. 1–6).
25. Cheng, N., Zhang, N., Lu, N., Shen, X., Mark, J. W., & Liu, F. (2014, janv.) Opportunistic spectrum access for CR-VANETs: A game-theoretic approach. *IEEE Transactions on Vehicular Technology, 63*(1), 237–251.
26. Perez-Romero, J., Sallent, O., Ahmadi, H., & Macaluso, I. (2016, avr.). On modeling channel selection in LTE-U as a repeated game. In *2016 IEEE WCNC*, Doha, Qatar (pp. 1–6).
27. Sohn, I., & Lee, S. H. (2015, October). Distributed scheduling for coexistence of IoT wireless devices. In *2015, International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, South Korea (pp. 680–682).
28. Yu, H., Iosifidis, G., Huang, J., & Tassiulas, L. (2016). Auction-based coopetition between LTE unlicensed and Wi-Fi.
29. Hamidouche, K., Saad, W., & Debbah, M. (2016). A multi-game framework for harmonized LTE-U and WiFi coexistence over unlicensed bands.
30. Do-Yun, K., & Choi, Y.-J. (2015). Spectrum coordination of multi-RAT D2D for IoT service. In *2015 ICTC*, Jeju Island.

Phishing Website Detection Using Machine Learning

Prateek Gupta and Archana Singh

Abstract Phishing costs around billions of dollars per year to the Internet users. Novel phishing techniques for instance spoofing in between trusted websites on the Internet are leveraged to phish target's account information, login credentials and personally identifiable information such as email Id, date of birth, biometrics and passwords. Most commonly attackers use phishing software and spam emails for stealing personal information to collect financial accounts details like credit card details and credentials. In the recent study, it was identified that phishing attacks account for more than 80% of reported security incidents and 94% of them are via email. The research work presented here is on a multitude of strategies which are used for detection of malicious and phishing sites depending on their various lexical features. Information gathered from the study of malicious and phishing sites is then used for lexical features assessment, and further, to analyze and to improve upon the algorithm used for the detection of malicious and phishing sites. Most organizations today use rule-based engines for phishing detection that do not proactively scale for phishing attacks without additional rules deployment. From the recent study, one could gather that for an organization, an improvement in phishing detection does have a positive impact on net revenue. An associated data point here is that $17,700 is lost every minute due to phishing attacks, thus a need for a comprehensive solution to phishing attacks. The paper provides an analysis of various methods used for detecting phishing websites by using machine learning and classification techniques based on lexical features. Machine learning-based techniques leverage natural language processing and other classification techniques like logistic regression, support vector machines and random forest [1, 2]. In order to have a comprehensive machine learning-based solution, training data is required to possess lots of relevant and non-correlated features. A comprehensive learning algorithm can effectively determine not previously classified URLs with a better accuracy. Here,

P. Gupta (✉)
Adobe, Ghaziabad, India
e-mail: pratgup@adobe.com

A. Singh
ASET, Amity University, Noida, UP, India
e-mail: asingh27@amity.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_19

our purpose is to study various machine learning algorithms that can be leveraged for safeguarding users from spoofed websites and help them not fall in the trap of phishing by detecting these websites early. Previous work done on the subject is also studied and compared against for accuracy.

Keywords Logistic regression · Random forest · Support vector machine · Security · PII · Phishing · Malicious

1 Introduction

The conventional method used to detect malicious websites is based on a predefined dictionary of blacklisted websites. A website gets blacklisted based on user's feedback who encountered a malicious intent. However, the disadvantage of the blacklisting method is that the websites which are not found previously as malicious cannot be predicted whether they are malicious or not. Blacklisting is a static method of detecting URLs malicious nature. Blacklisting method is mostly accurate and reliable. Nonetheless, it cannot be the only way of detecting malicious sites as the blacklisting method would not work accurately in today's time of dynamic malicious URLs. Dynamic here refers to auto-generating malicious URLs. As can be seen from the latest APWG phishing activity trends report, the number of phishing websites is significantly increasing year on year. Magnitude of increase can further be envisaged by the fact that there are many which get detected, blacklisted and closed—despite which graph is having an upward trend (Fig. 1).

Another disadvantage of the blacklisting method is that it is highly dependent on the incidents reported by the users. The heuristic classification is an improvement of the previously mentioned blacklisting approach. In this approach, the signature of the previously existing malicious URL and the signature of the new URL are matched. Even though a heuristic classification method like the blacklisting method is also highly effective, it cannot cope up with today's evolving phishing attack techniques [3]. Another demerit to consider is that both heuristic and blacklisting methods become more and more complex as the database of signatures, and blacklist

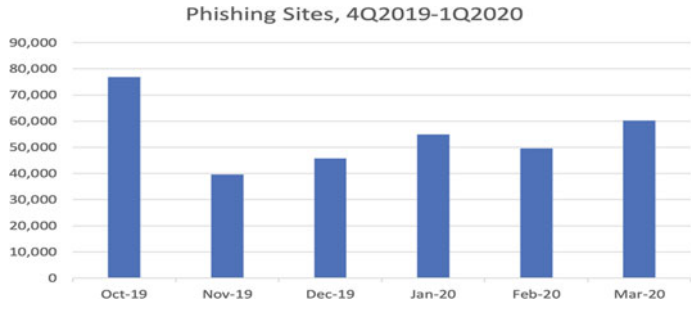


Fig. 1 Phishing sites from quarter year 2019–2020 (*source* Internet)

Fig. 2 Most targeted phishing sectors 2019–2020 (*source* Internet)

websites keep on increasing exponentially daily and at times even hourly. Phishing is not restricted to a sector and is well distributed. Each sector has its own set of phishing experts who deploy unique methods and semantics-based attack structure, which is designed specifically for that domain and mainly exploits the computer users' vulnerabilities (Fig. 2).

2 Features of URLs

2.1 Blacklist Features

As we have already discussed in the previous sections that the conventional method used to detect malicious websites is to make a list of blacklisted malicious websites. These blacklists are built using various methods such as feedback taken from humans whenever they encounter these websites, and the human feedbacks are very accurate in nature as they are verified by the humans itself. But the disadvantage of this blacklisting method is that the websites which are not found previously as malicious cannot be predicted whether they are malicious or not. As mentioned above, this

method is highly dependent on the incidents reported by the users. But it plays a vital role while the training of a machine learning algorithm, and hence, is considered in the feature list.

2.2 Lexical Features

The lexical features are also known as URL-based features. Foremostly, URL of the website is analyzed in order to detect malicious websites. In this length of the URL is taken into consideration, the number of digits in a URL is counted and typo squatted URLs are scanned (e.g. www.goggle.com). In addition, the number of sub-domains in the URL and whether the top-level domain is commonly used or not is also examined. Further, the algorithm is advanced to consider the number of dots in a URL and to dynamically identify domain and sub-domains.

Lexical features also examine tokens in hostname. To list a few: '?', '+', '%', '=', '.', etc. These features are helpful in verifying behavior of a webpage. For instance, multitude of tokens attributed to slashes may indicate denial of service attacks. The domain name can directly indicate a malicious website which has been previously blacklisted.

2.3 Features: Examine IP Address of the URL in Address Bar

For this feature, let us leverage the IP address of the given domain as alias, for example: "http://128.68.1.94/fakewebpage.html". To analyze cases wherein end users are sure that a person or a bot is seeking and trying to perform data theft on their personal information, IP address is sometimes even converted into the hexadecimal formatted code, as shown here: "http://0x58.0xCC.0xCA.0x62/2/paypal.ca/index.html".

$$\underline{Rule} : \text{IF} \begin{cases} \text{If The Domain Part has an IP Address} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Lengthy URL conceals the malicious element
Usually, what phishers do is, they make use of lengthy URLs, to make sure that the suspicious part in the URL is not visible. To strengthen our work and study, we figured out the size of URL in the given tuples of our dataset and tried to harness a basic URL length. The outcome and result reflected that phishing URLs mostly have the length of the URL similar to or greater than 54 alphabets and characters. Outcome was strengthened by studying the dataset we used wherein we figured that suspicious URL sizes are actually greater than 54 alphabets.

$$\text{Rule: IF}\begin{cases}\text{URL length} < 54 \rightarrow \text{feature} = \text{Legitimate}\\ \text{else if URL length} \geq 54 \text{ and } \leq 75 \rightarrow \text{feature} = \text{Suspicious}\\ \text{otherwise} \rightarrow \text{feature} = \text{Phishing}\end{cases}$$

We were in the position to be able to reconsider and update this particular attribute or feature, with the help of using a technique, which is totally based upon the frequency, and this helped us in increasing the relative accuracy of the algorithm.

Very Short URLS or "TinyURL" Using URL trimming services
Trimming services used by many phishers shorten the URL. It is a very smart method on the Internet, in which a URL which is relatively smaller in size but can cause equally or more detrimental attacks. This can be obtained and done, with the help of "HTTP. Redirect" upon a URL, which is small, that can be used to link toward the website, which has a very lengthy URL name. For example:

$$\underline{Rule} : \text{IF}\begin{cases}\text{TinyURL} \rightarrow \text{Phishing}\\ \text{Otherwise} \rightarrow \text{Legitimate}\end{cases}$$

URL's having "@" Symbol
When a phisher uses the "@" character in the URL, it results in the leading of the specific web browser to ignore values preceding the "@" symbol.

$$\text{Rule: IF}\begin{cases}\text{Url Having@Symbol} \rightarrow \text{Phishing}\\ \text{Otherwise} \rightarrow \text{Legitimate}\end{cases}$$

Redirecting using "//"
Existence of character "//" in the given URL path implies that the end user on click of URL will be sent to a non-identical webpage on the Internet. An example of such a website URL is "http://www.legitimate.com// http://www.phishing.com" [3, 4].

During our research, we examined a similar URL where there is a presence of "//". We identified that, if a URL starts along "HTTP", this implies that "//" will mostly appear in the sixth position of the URL. However, in a scenario wherein Domain URL has "HTTPS", then "//" will most likely be at position 7.

$$\text{Rule: IF}\begin{cases}\text{The Position of the Last Occurrence of "/" in the URL} > 7 \rightarrow \text{Phishing}\\ \text{Otherwise} \rightarrow \text{Legitimate}\end{cases}$$

Using (−) to the domain
This hyphen of dashed symbol is very rare in use in the good URLs. Attackers tend to attach prefixes and suffixes, which are separated by using the symbol of (−) in the domain name, in result of that, the end users tend to feel that he/she is handling the legitimate URL. For example:

$$\text{Rule: IF} \begin{cases} \text{Domain Name Part Includes}(-)\text{Symbol} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Sub-domain and Multi-sub-domains
If in case there is a situation, in which the number of given dots in the URL is much larger in terms of the number than 2 or 3, then we will identify such URL as "suspicious", since it has only one and not more than one sub-domain.

$$\text{Rule: IF} \begin{cases} \text{Dots In Domain Part} = 1 \rightarrow \text{Legitimate} \\ \text{Dots In Domain Part} = 2 \rightarrow \text{Suspicious} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

HTTPS (Hypertext transfer Protocol secured using Socket layer/TLS)
HTTPS in URL is crucial and helps in giving the impression of the webpage accuracy and reliability, but this technique cannot be used in isolation as it is not adequate enough. The writers of different research papers have given suggestions like to verify the assigned legitimate certification of the HTTPS which is to verify with the issuers if the domain is certified and the age of certification is valid. Certification authorities can be consistently listed among out most of the top renowned names like "geotrust, GoDaddy…, etc.". On further research and testing the dataset, we were able to find out that the minimum age of the given URL is of two and more years [5, 6].

$$\text{Rule: IF} \begin{cases} \text{Use https and Issuer Is Trusted and Age of Certificate} \\ \quad \geq 1 \text{ Years} \rightarrow \text{Legitimate} \\ \text{Using https and Issuer Is Not Trusted} \rightarrow \text{Suspicious} \\ \text{Otherwise} \rightarrow \text{Phishing} \end{cases}$$

Domain Registration Length
On the basis of the various past studies, it can be concluded that the phishing webpage lives only for a short period of time, while on the other hand legitimate domain names are generally paid for a longer duration, say for multiple years. During the analysis of the dataset in consideration, it was identified that the big and longest fraud domains are most likely to be used for not more than a year time frame.

$$\text{Rule: IF} \begin{cases} \text{Domains Expires on} \leq 1 \text{ years} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Favicon
Favicon is a picture associated with the specific website. Many users, which exist on the Internet, acting as an agent, such as graphic-oriented browsers and news apps,

leverage favicon to optically remind of the webpage and to add falcon as URL's identity in the address bar. In scenarios wherein the favicon is reloaded on refresh of URL and it apparently is different to the one shown in the address bar, then in that case, the website is considered as a malicious and a phishing website.

$$\text{Rule: IF} \begin{cases} \text{Favicon Loaded From External Domain} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Using Non-Standard Port

Non-standard port feature is extremely beneficial in validating if a service is from the authentic server. To prevent organization from phishing attacks, it is good to only open and close the ports as per requirement and keep them always up to date. Multiple security firewalls, proxy and NAT servers are placed to protect confidential data and to block the susceptible ports [7, 8].

$$\text{Rule: IF} \begin{cases} \text{Port \# is of the Preffered Status} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Using "HTTPS" Token in the domain element of the URL

The attackers can try to add the "HTTPS" token in the domain element of the URL, in order to disorient the users. For example:

$$\text{Rule: IF} \begin{cases} \text{Using HTTP Token in Domain Part of The URL} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Abnormal-based Features Request URL

It is validated if the different aspects in a webpage like image, visuals, graphics, videos and sounds are being directed to another domain. In a legitimate webpages, the webpages, and the content of webpage are mostly unique and have correlation only with other webpages in the same domain [5, 9, 10].

$$\text{Rule: IF} \begin{cases} \%\text{ of Request URL} < 22\% \rightarrow \text{ Legitimate} \\ \%\text{ of Request URL} \geq 22\% \text{ and } 61\% \rightarrow \text{Suspicious} \\ \text{Otherwise} \rightarrow \text{feature} = \text{Phishing} \end{cases}$$

URL of Anchor

An anchor URL is defined using the <a> tag. It is also known as link label and is examined to identify the rankings that the webpage will receive from globally renowned search engines. This feature is used as a "Request URL". Moreover, in this attribute, we validate:

Whether the <a> tag is partially or entirely associated with webpages and if all are corresponding to the same domain and have a similar ranking across search engines.

$$\underline{Rule} : \text{IF} \begin{cases} \%\text{ of URL Of Anchor} < 31\% \rightarrow \text{Legitimate} \\ \%\text{ of URL Of Anchor} \geq 31\% \text{ And } \leq 67\% \rightarrow \text{Suspicious} \\ \text{Otherwise} \rightarrow \text{Phishing} \end{cases}$$

Linked webpages are placed with <Meta>, <Script> and <Link> tags During the course of research, multiple methods that can be used in a webpage source code were covered. It was observed that it is acceptable for a website to make use of <Meta>, <Script> and <Link> tags to extract other resources.

$$\text{IF} \begin{cases} \%\text{ of Links in ``<Meta>'', ``<Script>'' and ``<Link>''} < 17\% \rightarrow \text{Legitimate} \\ \%\text{ of Links in ``<Meta>'', ``<Script>'' and ``<Link>''} \geq 17\% \text{ And } \leq 81\% \rightarrow \text{Suspicious} \\ \text{Otherwise} \rightarrow \text{Phishing} \end{cases}$$

Server from Handler (SFH) SFHs that contain an empty string and a "about.blank" are considered as suspicious. In addition, if a webpage's domain name is in the SFHs and is non-identical to the parent domain name, it implies a suspicious website.

$$\text{Rule: IF} \begin{cases} \text{SFH is ``about:blank'' Or Is Empty} \rightarrow \text{Phishing} \\ \text{SFH Refers To A Different Domain} \rightarrow \text{Suspicious} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Submitting information to Email Internet forms, which are used to fill the information, require a user to provide his personally identifiable information. These Internet forms are directed to a backed computer known as a server for further dealing and processing. An attacker may redirect this data and information to his own computer and his workspace and his storage. In the subsequent steps, a script running on the server side of the connection can be accessed for his personal use.

$$\text{Rule: IF} \begin{cases} \text{Using ``mail()'' or ``mailto:'' Function to Submit User Information} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

Abnormal URL

This kind of attribute can be taken out of the WHOIS database, available as a library in Python.

$$\text{Rule: IF} \begin{cases} \text{The Host Name Is Not Included In URL} \rightarrow \text{Phishing} \\ \text{Otherwise} \rightarrow \text{Legitimate} \end{cases}$$

3 Conclusion and Future Scope

As we can see in this paper, malicious URLs are a big problem in today's world. There has been prior research done in this field, and we have added our findings to the same. However, our approach was to see all aspects related to this field. From our research, we understood that many attempts have been made for the detection and prevention of malicious and phishing URLs present on the web. On reading different research papers and publications, we saw that machine learning is the common approach, which is being ardently followed by people who are working in this area. As we know machine learning is vast in its reach and the impact that it can create, so usage of machine learning for detection and prevention from malicious URLs is good and beneficial. As we identified, a URL is made up of different parts, URLs are just not like a name of a person and have many of their own characteristics. URLs are made up of different elements like its domain, sub-domain, port address, etc. Therefore, what machine learning models do is that it learns tons of URLs and tries to find the similarity between them. So, URLs having the same type of structure are classified as a particular class. Therefore, if we give our model the data of malicious URLs, it will go through every one of them. Then it will try to find similarity between them and classify them as a class of URLs. So when we give any URL which is of the same structure as malicious URLs, a machine learning model will identify that this is a phishing or a malicious URL. So, in machine learning, different algorithms and techniques are present to train a classification model, and each algorithm has its own benefits and disadvantages. Each of them has a different way of learning.

In our future work, the work can be extended by considering more features of URL phishing, and analysis can be done using deep learning techniques and transfer learning. The work can be extended for the dark web to provide more secure solutions.

References

1. Sahingoz, O. K., et al. (2019). Machine learning based phishing detection from URLs. *Expert Systems with Applications, 117,* 345–357.
2. Jain, A. K., & Gupta, B. B. (2018). PHISH-SAFE: URL features-based phishing detection system using machine learning. In M. Bokhari, N. Agrawal, & D. Saini (Eds.), *Cyber Security. Advances in Intelligent Systems and Computing* (Vol. 729). Singapore: Springer.
3. Buber, E., Dırı, B., & Sahingoz, O. K. (2017). Detecting phishing attacks from URL by using NLP techniques. In *2017 International Conference on Computer Science and Engineering (UBMK)*, Antalya (pp. 337–342). https://doi.org/10.1109/ubmk.2017.8093406.
4. APWG. (2020, December). Phishing activity trends report, 1st Quarter 2020. Technical Report.
5. Vargas, J., Correa Bahnsen, A., Villegas, S., & Ingevaldson, D. (2016). Knowing your enemies: Leveraging data analysis to expose phishing patterns against a major US financial institution. In *2016 APWG Symposium on Electronic Crime Research (eCrime)* (pp. 52–61).
6. Ma, J., Saul, L. K., Savage, S., & Voelker, G. M. (2011, May). Learning to detect malicious urls. *ACM Transactions on Intelligent System Technology, 2*(3), 30:1–30:24; Marchal, S., Saari, K., Singh, N., & Asokan, N. (2016). Know your phish: Novel techniques for detecting phishing sites and their targets. In *International Conference on Distributed Computing Systems* (pp. 323–333).

7. Woodbridge, J., Anderson, H. S., Ahuja, A., & Grant, D. (2016, November). Predicting domain generation algorithms with long short-term memory networks. http://arxiv.org/abs/1611.00791.
8. Zhang, J., Porras, P., & Ullrich, J. (2008). Highly predictive blacklisting. In *17th USENIX Security Symposium* (pp. 107–122).
9. Roopak, S., & Thomas, T. (2014). A novel phishing page detection mechanism using html source code comparison and cosine similarity. In *2014 Fourth International Conference on Advances in Computing and Communications* (pp. 167–170).
10. Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. In *SIGCHI Conference on Human Factors in Computing Systems* (pp. 581–590).

Engineering in Action

Cat Kutay

Abstract Engineering education is invariably highly technical and leaves little time for experience of the vagaries of real life. While industrial experience is now an established part of most professional degrees, this experience is usually limited to large business applications that are selected in areas of the student's knowledge. As a result, professional courses are producing engineers that lack the experience to carry out project work, to communicate effectively with the user in a way that produces sustainable products. Can we claim to have produced professionals? This paper looks at two aspects of teaching sustainability and cross-cultural skills to professionals: firstly, the need for students to understand how technology is related to its cultural and environmental context, how it must be designed within this context and existing technologies may not be transferable from that context and, secondly, the need for students to see their role as educating the public as well as themselves. By using examples from projects in remote communities in Australia and the Pacific, this paper shows that much of the learning students will otherwise gain throughout their lifetime can be encapsulated in small projects during their undergraduate years. Software tools can be developed to present this wealth of information to learners at all levels.

Keywords Sustainable engineering education · Community engagement · Ethical engineering

1 Introduction

In the education of any professional, the ethical understanding achieved by students is important. In the present state of climate change and need for innovation around sustainable production, the teaching of sustainable practice has become an ethical consideration. However, it is not easy to teach these principles in a way that students can engage. We discuss ways to implement this imperative.

C. Kutay (✉)
Charles Darwin University, Ellengowan Drive, Casuarina, NT 0810, Australia
e-mail: cat.kutay@cdu.edu.au

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_20

This paper extends the concept of professional training in three ways. Firstly, we need to cover responsibility for the effect of designs and construction recognising sustainability as being a fundamental part of any design. Secondly, these processes can be taught to our students so that they have to tools to undertake sustainable projects, rather than simply the moral desire to do so. The final section of this paper looks at engineers as trainers and how this role is important for ethical and sustainability considerations in the projects discussed. We aim to train students to consider themselves as part of a sustainable community.

2 Ethical Engineering—Sustainability in Engineering

Sustainability is defined as 'development that meets the needs of the present without compromising the ability of future generations to meet their own need', in the Brundtland Commission Report (1987). This attribute of projects is hard to teach as we have limited knowledge of the dynamic interactions between the three pillars in any systems, particularly the inertia and the strength of interactions.

In the small projects Sect. 4 below, we look at projects which tackle issues of sustainability. The vulnerability or resilience of the environment-social-technology system varies between different places and for types of ecosystems and human livelihoods. This can be seen in adapting projects from one remote settlement to another in Australia in Sect. 4.1.

The long-term aim is to provide scientifically meaningful measures that would alert when conditions go beyond where environment-social-technology systems incur a significantly increased risk of degradation. For a simple example, we look at the Indonesian PV feasibility study conducted by a student in Sect. 4.2.

Long-term trends in environment and development, including local consumption and population growth, shape environment-social-technology interactions and were used as the basis for the micro-hydro projects in the Solomon Islands in Sect. 4.3.

We are integrating our present relatively independent activities of planning, monitoring, assessment and decision support into systems for adaptive management and societal learning. This is dealt with in engineers as trainers, Sect. 5.

3 Ethical Engineering—Using an Engineering Process

In engineering industrial projects, the actual process is rarely visible. Most 'real' projects are of too long a duration for students to experience a significant proportion during their working period. This suggests the need to focus students experience on small projects or at least on stages in the project that show significant transitions, for instance: project initialisation, design and evaluation.

The notions of reflection-in-action and reflection-on-action were central to Schön's [12] contributions in this area of professional training. The first is described

as ‘thinking on our feet’. It involves using our past experiences, our present feelings and using our theories to apply to the present situation. This involves building new understandings to inform our actions in the situation that is unfolding, based on our prior reflection-on-action.

Kolb [8] also modelled an approach to learning as a cycle based on experience and reflection. He stressed that the process is more important than the outcome in learning and that it is a social process. Kolb defined learning as the process through which knowledge is created by the transformation of our experience [8].

Schön was concerned with the technical rationality view of knowledge which forms a positivist epistemology of practice and leaves a dichotomy between rigour and relevance for professionals to confront. He suggested an alternative epistemology ‘in which the knowledge inherent in practice is to be understood as artful doing’ [13]. Learning contexts that provide student with good flow, or immersion, tend to motivate further involvement and hence opportunities to learn [12].

Argyris and Schön [1] proposed that professionals have mental maps on how to act in a given situation. This is used in planning, implementing and reviewing their actions. Furthermore, they asserted that people use these maps to guide their actions rather than theories, which are used more for post-explanation.

Another point that is relevant to the present research is how Schön talks of training. Students often learn in a practicum, an environment that resembles the real world, with peers and a tutor/mentor to interact with. Their background learning, and their experience by doing, also assists in the learning. Schön sees the role of the coach in this situation as involving demonstration and self-description as well as comparing process and product, for example ‘did the process work?’. The students have to devise new methods of reasoning for reflection-in-action and also construct and test new mental models and ways of framing problems to recognise the familiar in future.

4 Small Projects—Community Projects for on-Site Learning

This paper proposes that the lack of complexity in engineering models for teaching does not provide students with the experience or skills needed to analyse real projects. The following are projects where students are exposed to the complexity of issues involved in project management and communication.

In these projects, it is important that students have good support on the project, or the difficulties will override the benefits they gain from such experience. Also, the projects themselves are still large, encompassing many years work for a single community in terms of fundraising, preparation and construction work, besides the installation and maintenance training. Within these, smaller projects such as installation of solar water heaters, pit toilets, pour flush and low flush toilets, monitoring of renewable systems, radio training, etc. could be handled by a single student project.

Even involvement in a short training programme on-site for turnkey solutions is constructive experience [15].

4.1 Australian Aboriginal Communities

There are a large number of Aboriginal communities in Australia that lie outside town council boundaries. These include large settlements and outstations. Their isolation from mainstream services including basic maintenance has provided ample opportunity for engineers and tradespeople to design and apply alternative technologies.

From years of supervising students on such design and installation projects, the main lesson we learnt was that such projects are each unique. In trying to provide information to students on previous projects, we always had to stress they would still have to learn about the new situation, whether it is a new technology or a new community group involved. However, this is possibly not a fixed feature of Indigenous technology projects. For instance, the Navajo centralised the development of their Internet access [9] as does the First Nations Media centralise their collection and distribution of media [5].

There were perhaps five major factors that made these project significant learning experiences:

- **Isolation**. Instead of being able to refer to texts or experts, students had to solve the problem on-site. Even tradespeople made mistakes in standard procedures here, such as installing septic tanks reversed so that the inlet is below the overflow.
- **Organisational Culture**. The organisational methods in Aboriginal communities are different to European and can differ between communities. When tools are removed from the community shed, any equipment that needs repairing is likely to be found with the tools, rather than the tool returned to the shed to carry out the repairs. This is a process that can be worked with, rather than fought.
- **Understanding how technology would be used**. Observation skills have to be developed to assess how the community will work with the technology. Sewerage evaporation ponds are pools. Children in a large settlement were swimming in the third sewerage evaporation pond, as they had been told it was 'clean' by the 'waterman'.
- **Explaining the technology to the community**. The language and conceptual barriers were vast. For instance, the notion of cause and effect may be different as seen in the previous example where the children said the water did not make them sick, but they did get ear and throat infections.
- **Involving the community in maintenance**. The only way any of these technologies could survive was to involve the community from the project inception, such as it was necessary for students to provide training during installation.

We developed problem-based courses that encouraged locals to bring up their own issues from the community in the specific domain areas of the course (including infrastructure, housing, communication and energy).

The issues of infrastructure and technology need for these communities remains. While much technology has been developed, it still needs to be tailored to each community and the instruction provided on-site for each system. These projects provide industrial style experience that requires two on-site visits with intervening period for system design and producing teaching materials such as posters.

4.2 Indonesia Solar Electricity

Installation projects in photovoltaics (solar electricity) can be small and carried out by a student group. However, the value to community has to be assessed. The issues encountered on these projects included:

- **Training Photovoltaics (PV)** is still new to many communities, so its use can entail a cultural gap between the system and the users which often leads to poor performance or a shorter life for PV systems [10] as this is not a familiar technology.
- **Using the technology**. A follow-up visit to an installed system revealed that many users found that the PV lamp was too bright when sleeping and could also attract the attention of thieves. Consequently, people still used traditional kerosene lamps after 10 pm, adding the cost of the PV system which they had paid to install, to their energy bill.
- **Finance and Maintenance**. Adequate transfer of knowledge must also occur at the administration level to ensure availability of trained technicians and spare parts. System outages encouraged non-payment from the users which became contagious.
- **Feasibility**. In order to verify the environmental sustainability of PV, students need to estimate the reduction in CO_2 emissions and other environmental and health costs beyond just the cost savings.
- **Risk Management**. As a related aspect of the project, we need to consider the disposal of PV system wastes such as batteries, lamp tubes and other electronic components; how to inform the villagers about proper waste handling and disposal; and the effects on human health of the waste.

4.3 Micro-hydro in the Solomon Islands

Micro-hydro systems have been installed in remote communities in the Solomon Islands to provide an economic base for villages separate from the logging and sawmills that dominate rural areas. The isolated nature of the communities provided greater opportunity to observe the support needed on technology projects.

- **Management**. In the Solomon Islands, it has been observed that there are a number of half-finished micro-hydro plants. These projects [3] show the importance of community management and ownership and setting up such structures is often lengthy and complex. The process is more than a construction project and needs strong local participation.
- **Cultural Processes**. The process of maintenance and control of a new technology may not necessarily fit easily within the exiting community management infrastructure and can create tensions. While this should be accessed before the technology is introduced, sometimes the community may decide the benefits outweigh the difficulties.
- **Centralisation**. There has emerged a clear recognition by each community in the Solomon Islands involved in these projects, of the need for institutions to 'collectivise' their isolated and individually controlled power systems, in the crucial areas of skills, spare parts, manpower and economic applications for the electricity.

5 Engineers as Trainers—The Learning Society

Schön promoted the vision of a learning society [12], where all people are continually learning. In sustainable and remote development projects, this is crucial. In a report from a survey of electrical infrastructure in Papua New Guinea [7], it was noted that electrical accidents were a problem which is not limited to remote communities. Not only must engineers today know more about the ways that the public uses and misuses technical artefacts, but also the highly technical basis of modern society means that familiarity with technology needs to be a growing part of higher education for everyone, not just for engineers [14].

We note that all the issues that arose in the above project examples still apply in mainstream projects run by large firms in Australia. However, the latter provides less examples of situations that are 'beyond the text', so require little awareness of the process involved. In the area of consumer understanding of the technology, this is most evident. While peoples' understanding of micro-hydro will be limited by their previous experience with outboards or generators in the Solomon Islands, these are plenty of examples where information is not provided to the consumer, either deliberately or through neglect.

Another aspect that arises in community projects is the problem of using a simple model to describe a system which can fail to apply as knowledge of the system increases. For instance, in explaining the use of solar panels to recharge batteries in the Solomon Islands, we used a simple analogy with the use of the sun to supply energy (heat) to our bodies. Users were considering putting their old batteries in the fire 'to recharge them'.

This suggests a flexible training format where the depth at which the learners wish to approach the subject can be selected and the course can re-design around these parameters. Such systems are being built based on learning ontologies that define the learning role and level of various documents and resources for learning. The learner

can then trace and re-trace the course format at varying levels of sophistication. It is important that a range of community-based projects contribute to the knowledge base of these learning systems.

5.1 Learning Process

The process we used in this training is:

1. Provide material on the technology to be used in the field.
2. Students develop a standard system design. Much of the equipment can be organised for the installation at this point, as the technology variation is limited.
3. Provide an environment where the functioning of a community is examined. We run pre-workshops where we provide scenarios where examples of real technology installations have been carried out. This is followed up during any later field trip where they are required to problem solve in real situation.
4. Provide an environment where assumptions about the functioning of the system they have designed are questioned. If possible, students take initial field trip to community to meet the potential users, to survey the needs or to analyse existing similar installations.
5. Student report on the final system installation and maintenance process as proposed, including any training material that may be required.
6. Students fundraise and travel to location to carry out installation or training of maintenance team.

5.2 Tools for Learning

There is a system developed to author courses in renewable energy based on XML documents and developed around a learning ontology for collaborative learning (CL) proposed [2] which is centred on activity theory. The challenging part is to virtualise the CL scenario, by constructing intelligent user profiles, in which project profiles could be properly reflected. Isotani et al. [6] provide a framework or model to describe the content and learning level of documents and to create a structure for instantiation according to the learners' requirements or in this case that of the community project.

The tool includes an editor to set up documents for viewing by students. Then documents can be combined by an instructor to provide a link between various resources on a subject. The system will require that an instructor set up the framework for the course, but the learner interface will then function independently. Students can revisit the material over time as their understanding grows.

A constructive learning approach is taken in these courses which means that the students are encouraged to construct their understanding from the information available [11]. Students can also use the tool to edit their own documents according to different views or interfaces to a report. This is designed to help them analyse their

work from different level and also ensures documents can be included in a knowledge repository on that domain.

While most tools to support course development focus on supporting the instructor in developing a suitable interface, this tool is designed to focus on the instructor developing a suitable knowledge base for the students to manipulate. The task of designing pedagogically sound CL scenarios requires a deep understanding of instructional theories and skills to organise and refine the CL scenario. Some of this approach is being used to develop the training material for the EWB Challenges in Australia on an interactive LearningX site [4].

6 Issues for the Engineers

During the running of these projects, the features that are most striking are:

- The students learn from their mistakes. While they do not have final say in decisions, their recommendations are considered and hence must be thorough. This provides opportunity for reflection-on-action and re-thinking their concepts and assumptions about the technology and the users.
- The community learns as the students learn. While students are re-thinking their ideas, they are working at a more basic level with the concepts. This makes them more able to answer the fundamental questions put to them either by their observations or by community members. If they have not re-thought significant aspects of the project, the community of users will ensure they do.
- The skills that students develop do not apply only to exotic and remote projects. They learn skills that apply in any engineering project, although not usually in such a clear manner, as the difference in community approach will be less obvious.

The skills relevant to an engineering or management career that students gain include aspects of the project management process and those related to communication with the user. There are more than discussed here, but they provide some examples.

- **Project Management**. In communities, people tend to have much clearer defined roles in relation to the project as these have to be created or adapted from old management processes. The students are at all time aware of their role in the community and have to question this. Also, the scope of the work is clarified early, so as not to influence community life beyond this. The project forms a strong team as it relies heavily on local labour, firstly as clients for the training and secondly as the cost is reduced by relying on human labour rather than expensive installation work. Also tools or local equipment may be adapted, such as construction of powerhouses.
- **Communication with the User**. Designing for the user in any new system or installation makes it clear how the system has been adapted to the user or the technology re-designed. This promotes students to think how even standard technology

is adaptable to different users. The projects are necessarily short, so students are aware of the need to hand over control of the technology from the start, but here it is not automatic, for instance 'here is a light bulb, replace it if it blows' is not so simple. Knowledge or prior learning needed in training to explain any technology to communities is clear as students have to find out what the people already know, what they use in their lives that may relate and how they may use the new technology.

7 Future Work

The limitations of this work revolve around concerns that the community will not be receiving high technical quality work, and time and money is being put into training students instead of community development. However, this time is often freely given by community as they value the role they play in the students' education.

The work also requires that mentors develop relationships with community which may not be possible or work with organisations like Engineers Without Borders [4] to engage their students in the EWB Challenge which involves such projects without any travel in the first year, but with the opportunity to visit communities in later years.

8 Conclusion

The present training of engineering professionals does not enable students to appreciate sustainability as an integral part of their projects. Secondly, they often do not require the students to acknowledge the assumptions used in conventional engineering to simplify the problems to the known conditions or to question any foundations of existing technology. When working outside the normal framework, these assumptions can be incorrect. Finally, sustainability must focus not just on the maintenance of technology through training, but also the control of technology by society through open learning. The skills that students learn in such projects are not limited in use to remote and low-technology applications. The nature of these projects however brings into stark contrast the processes that make any project successful and ones that will not.

References

1. Argyris, C., & Schön, D. (1974). *Theory in practice*. San Francisco: Jossey Bass.
2. Barros, B., Verdejo, M. F., Read T., & Mizoguchi, R. (2002). Applications of a collaborative learning ontology. In C. A. Coello, A. de Albornoz, L. E. Sucar, & O. C. Battistutti

(Eds.), *Proceedings MICAI 2002: Advances in Artificial Intelligence: Second Mexican International Conference on Artificial Intelligence Merida, Yucatan, Mexico, April 22–26*. Heidelberg: Springer.
3. Bryce, P., Bryce, D., & Irons, C. (1994, April 6–7). Rural electrification and technology transfer. In *Conference on Technology Transfer in Remote Communities*. Perth: Murdoch University.
4. EWB Challenge. Discover the impact of engineering through real-world sustainable development projects. https://ewbchallenge.org/. Last accessed 3/6/2020.
5. First Nations Media Australia. https://firstnationsmedia.org.au/. Last accessed 3/6/2020.
6. Isotani, S., Mizoguchi, R., Isotani, S., Capeli, O., Isotani, N., de Albuquerque, A., et al. (2013). A semantic web-based authoring tool to facilitate the planning of collaborative learning scenarios compliant with learning theories. *Computers & Education, 63,* 267–284.
7. Jansen, T. (2004). By foot to the interior-building self-reliance during the civil war in Bougainville. Available September 10, http://www.terracircle.org.au/reports/reports.html (2004).
8. Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.
9. Meer, J. (2003). Getting on the Net: Struggle for digital inclusion of the Navajo. *IEEE Technology and Society, 22*(1), 53–58.
10. Retnanestri, M., Outhred, H., & Healy, S. (2003). Off-grid photovoltaic applications in Indonesia: An analysis of preliminary fieldwork experience. In *ANZSES03*.
11. Savery, J. R, & Duffy, T. M. (1995, September–October). Problem based learning: an instructional model and its Constructivist framework. *Educational Technology, 355,* 31–38.
12. Schön, D. A. (1983). *The reflective practitioner. How professional think in action*. London: Kogan Page.
13. Schön, D. A. (1987). *Educating the reflective practitioner: Towards a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass.
14. Williams, R. H. (2002). *Retooling: A historian confronts technological change*. Cambridge, MA: M.I.T. Press.
15. Westerlund, L., Ho, G., Anda, M., Daniel, M., Wood, K., & Koshy, K. (2009). Case study of technology transfer to a Fiji rural village using an improved sustainable turnkey approach. In *Technologies and management for sustainable biosystems* (pp. 245–252). Nova Science Publishers.

Adaptive Tuning Parameters for Data Transmission over IEEE-802.15.4 Wireless Sensors Network

Imen Bouazzi and Monji Zaidi

Abstract Wireless sensors networks (WSNs) present the basic technology for development of such applications, as WSNs split the extra challenges of harsh environments and reduced power supply; those are the famous feature of the large use of this technology. The most difficult phase of WSNs is to provide a good quality of service (QoS) while sensing data. The problem of the medium sharing is solved by MAC protocol that provides an assumption to schedule data transmission. So, in this work, we propose a tuning strategy to dynamically perform data transmission according to the packet transmission delay and to prolong the lifetime of each sensor nodes by optimizing their energy consumption. A dynamic set of parameters for the method used to access the channel is conducted. This management is based on information about the queue length of sensor node in which a priority algorithm is made to order access to the channel according to the queue length situation. Simulation results of the proposed scheme show that this strategy performs well than the basic standard.

Keywords Packet transmission · Energy · Priority · Packet loss · Traffic

1 Introduction

WSNs become more and more important in our daily use. With the progress of technology, WSN has been used in many areas such as security, healthy and military. At present, this development faces many technical problems; the famous typical problems are the optimization of network performance. Numerous MAC protocols for wireless sensor networks adopt different ways of energy saving, such as the duty

I. Bouazzi (✉)
Department of Industrial Engineering, King Khalid University, Abha, Saudi Arabia
e-mail: izaidy@kku.edu.sa

I. Bouazzi · M. Zaidi
Electronic and Micro-Electronic Laboratory, FSM, University of Monastir, Monastir, Tunisia

M. Zaidi
Department of Electrical Engineering, King Khalid University, Abha, Saudi Arabia

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_21

cycling technique, packets transmission delay scheduling and traffic enhancement [1, 2]. These protocols are efficient in terms of energy optimization but not all of them are good for improving the QoS, especially when it comes to burst traffic. More recent work focuses on the development of MAC protocols allowing both a very low duty cycle when traffic is low or zero and a high speed (therefore low delay) when traffic becomes high. In this work, we analyze the most representative duty cycle MAC protocols to understand the fundamental principle of the design, which consists of minimizing over-listening, idle listening, additional protocol costs and collision. In this paper, we implement a mechanism that we recently developed; whose packet transmission delay self-adapts to the variation in traffic to enhance the QoS and to minimize the energy consumption. The remainder of this paper is structured as follow; Sect. 2 presents the IEEE-802.15.4-MAC standard, its superframe structure and the data transmission procedure. Then in Sect. 3, we discuss the implemented algorithm. Section 4 results present the simulation result of our implemented algorithm while Sect. 5 makes conclusion and future work.

2 IEEE-802.15.4 Standard

2.1 Overview

This standard, initially written in 2003, and then revised in 2006, 2011 and 2015, describes itself as defining "the compatible protocol and interconnection for data communication devices using low-frequency radio frequency transmissions of short-range, low-power and low-complexity data in a personal wireless network". It defines the characteristics of the physical layers, as well as the data link layer which manages these physical layers. The MAC layer has a role of organizing transmissions above the sub-adjacent physical layer. In the case of the IEEE-802.15.4 standard, this clearly gives us its scope: The functions of the MAC sub-layer are management of beacons, access to channels, management of guaranteed time slots (guaranteed time slots) and frame validation, delivery of acknowledged frames and association and dissociation. In addition, the MAC sub-layer provides interfaces (hooks) for the implementation of application-specific security mechanisms.

2.2 Transmission Mode

The IEEE-802.15.4 standard proposes several modes of communication; the famous method is the non-beacon-enabled mode, where hosts communicate with each other without time synchronization. This saves a lot of energy for the terminal nodes (which do not route the traffic of other nodes), but in return, this requires that the nodes in the network are constantly available in RX mode (reception mode for a

radio), but in this case, there is no energy saving for them. The transmission mode of the IEEE-802.15.4 standard proposes the use of beacon-enabled mode, allowing time synchronization of the devices. The emission of the beacon frame defines three time periods. The first, called period with access contention, allows any node to transmit to the transmitter of the beacon frame (the coordinator), but with contention management (among others), the nodes must listen before transmitting to avoid overwriting transmission. The nodes also use this period to receive messages that the coordinator has indicated, by explicitly requesting the sending of frames using the receiver-initiated transmission (RIT) mode.

3 Implementation

Data transmission will be scheduled according to the required time to send data packet. After several transmissions, we measure the delay requested to send one data packet. If these measured parameters have a constant value, it means that a periodic data transmission is scheduled, else if this value is a random one, this means that unexpected transmission can take place any time and node function cannot be predicted. The idea behind this management is to predict the sleep and active period duration of each node to optimize energy consumption in the network. The WSN, studied in this work, contains 30 nodes placed in a cluster tree manner. If the channel is sensed free, node i with high priority performs its data during CAP period and other normal node transmits their data during CFP period. We suppose that is no packet dropped, and the packet transmission delay is given as:

$$T_D[s] = \frac{(T_b + T_{\text{Packet}} + T_{\text{Round}} + T_{\text{ACK}} + T_{\text{IFS}})}{R_{\text{Data}}} \tag{1}$$

where $T_b, T_{\text{Packet}}, T_{\text{Round}}, T_{\text{ACK}}$ and T_{IFS} are the total backoff time, the time to transmit data packet, a turnaround time (T_x to Ack) symbols after the reception of the data frame (192 μs), time to transmit ACK packet and time duration of IFS, respectively as presented in Fig. 1.

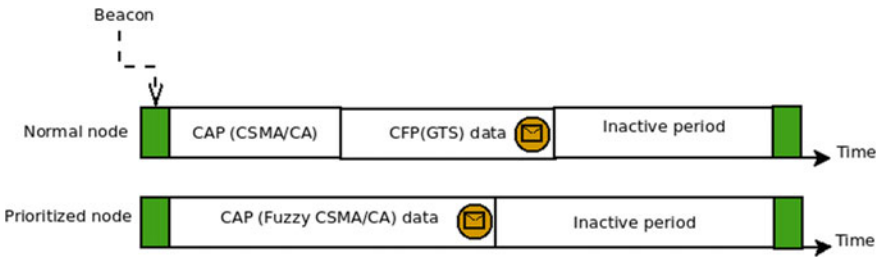


Fig. 1 New superframe structure

Suppose that every time a node i needs to transmit data, it should wait for a random BE period in the range of [0; 2^{BE}–1] before performing the CCA. Each time, the medium is sensed idle, and all nodes can send their data. Otherwise, if the senses channel is busy, nodes should wait for a random specific period before trying again to sense the occupancy of the channel. The needed access time that a node should wait before sensed the channel to be free is calculated as presented follow:

$$\begin{aligned} T_{\mathrm{CA}} &= \mathrm{InitialBackoffPeriod} + \mathrm{CCA} \\ &= (2^3 - 1) \times \mathrm{aUnitBackoffPeriod} + \mathrm{CCA} \\ &= 7 \times 320\,\mathrm{us} + 128\,\mathrm{us} = 2.37\,\mathrm{ms} \end{aligned} \tag{2}$$

In the CSMA/CA using the IEEE-802.15.4 [3, 4], the CCA detection time is defined as eight symbol periods. AUnitBackoffPeriod is defined as 20 symbol periods, where one symbol corresponds to 16ls. Figure 2 shows the packet scheduling delay, and we define then the total time needed to transmit one data packet in acknowledged transmission as follow:

$$T_t = T_D + T_{ca} + T_{\mathrm{BE}} + T_{\mathrm{ack}} \tag{3}$$

We have proposed in a previous work [5] a fuzzy logic algorithm that dynamically set three levels of priority for nodes to access the channel. This scheme improves the energy consumption by giving priority for nodes how to have a critical condition to access first to the channel. Due to queue size and traffic rate of each node, three degree of priority has been assigned for nodes (low, medium and high). The idea behind this work is to dynamically change the BE value to improve the network lifetime. Nodes with high priority should have the lowest value of BE to minimize the time to access in the medium. To more improve energy consumption, we have proposed an algorithm that classifies also nodes into tow category. Devices with high traffic send its data during the CAP period and nodes with a normal traffic transmit data during CFP period. This classification is accomplished according to the packet transmission delay. Whenever the measured TD parameter is a random value, the

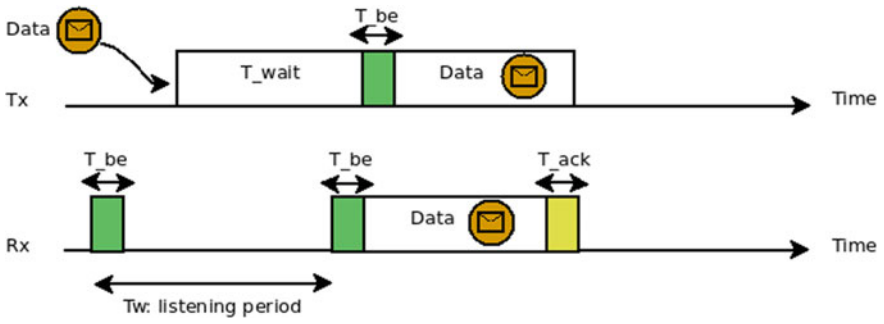


Fig. 2 Packet scheduling

transmission is performed during CAP period. When the required time to transmit data, packet has a periodic value; the transmission should be done during CFP period. Suppose that data transmission is done only during the CAP period using our fuzzy proposed scheme. The sleep period will be increased wish leads to improve energy consumption in the networks. In WSN, the power consumption in the active period is greater than the acquired power of a receiver node and greater than the power consumption in the sleep state, and we assume that

$$P_{tx} > P_a \gg P_{rx} > P_s \tag{4}$$

where P_{rx} presents the power consumed by receiver node, P_a defines the power wasted in during the active period, P_{tx} is the power harvested during the transmission, and P_s presents the power consumed during the sleep state. Let us assume that the power wasted during the active period is the same as the power wasted during the sleep period [6]:

$$E_a(t) + E_{tx}(t) = E_s(t) \tag{5}$$

E_a and E_{tx} define the amount of consumed energy during the active period [7]. E_s is the power wasted during the sleep period. Here, the relation between the powers can be written as follow (6):

$$E_a(t) = P_a T_a(t) - P_{rx} T_a(t) \tag{6}$$

$$E_{tx}(t) = P_{tx} T_{tx}(t) - P_{rx} T_{tx}(t) \tag{7}$$

$$E_s(t) = P_{rx} T_s(t) - P_s T_s(t) \tag{8}$$

From this equation, we can have

$$P_a - P_{rx} T_c(t) + P_{tx} - P_{rx} T_{tx}(t) = P_{rx} - P_s T_s(t) \tag{9}$$

Therefore, the duration of the sleep state can be estimated as a linear equation presented below:

$$T_s(t) = a T_a(t) + b T_{tx}(t) \tag{10}$$

where

$$a = \frac{P_a - P_{rx}}{P_{rx} - P_s} \tag{11}$$

And

$$b = \frac{P_{tx} - P_{rx}}{P_{rx} - P_s} \tag{12}$$

4 Simulation Results

We implemented our solution using NS-2 tool. The simulation results obtained clearly show the improvement of the proposed algorithm compared to other works. The simulation parameters used in this work are the same one used in [1].

We have evaluated the network performance in three metrics of packet delivery ratio (PDR), throughput and Jain's index which presents the equality in terms of access to the channel of all the nodes in the network. The PDR is defined as follow:

$$\text{PDR} = \frac{\text{Total Packets Received}}{\text{Total Packets Generated} \times 100} \tag{13}$$

Figure 3 presents the PDR variation in the network with respect of BO. For low value of BO, we notice that the PDR increases due to the scheduling of short superframe without collision. Whenever BO increases, the number of packets generated will increase which leads to apparition of collision between nodes. We notice that the DAMT program maximizes the PDR instead of the standard who still almost stable which exhibits poor performance. Almost the majority of packet do not reach their destination, and the cause is that the number of packets generated between two

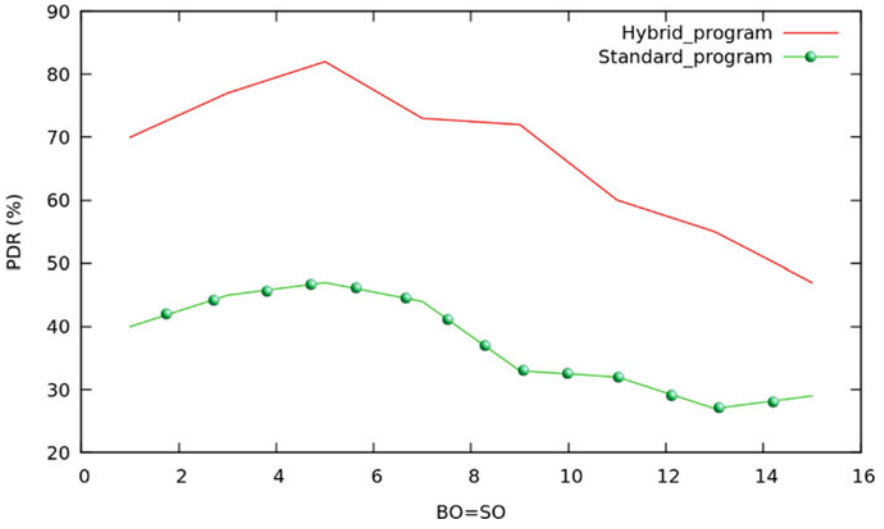


Fig. 3 PDR versus BO

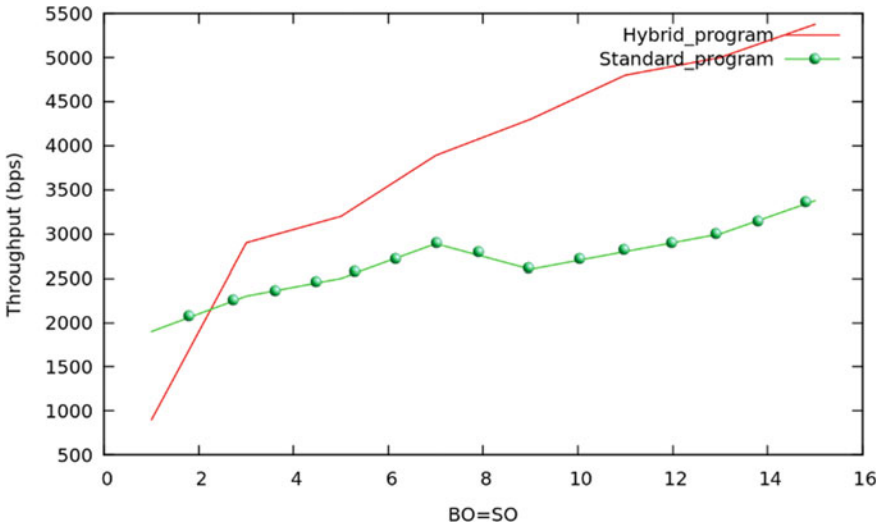


Fig. 4 Throughput versus BO

periods of activity is too large compared to the size of the active period, exceeding the capacity offered by the network.

Figure 4 presents the throughput variation under different value of BO and SO. We have fixed an equal value of the BO and the SO to see the impact of the superframe duration on the throughput. As seen in this figure, when the value of the beacon order BO increases, the throughput increases too, this means that more the length of the active period rise this May have an unused channel for long superframe duration. As shown for the red curve, our proposed algorithm is more significant than the standard IEEE802.15.4 in as long as he maximizes the throughput in the network.

To evaluate our heuristic approach, we have used Jain's fairness index to evaluate the fairness in the network. This metric is very important to improve the efficiency of an approach. Fairness of a protocol is its capability to assign a fair channel bandwidth as follow [8]:

$$F = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \times \sum_{i=1}^{n} x_i^2};\ 0 < F < 1 \tag{14}$$

$$\text{Here } x_i = \frac{\text{Ratio of actual throughput}}{\text{Optimal throughput}} \tag{15}$$

$$\text{Where } x_i = \left\{ \begin{array}{l} \frac{a_i}{b_i} \text{ if } a_i < b_i \\ 1 \quad \text{Otherwise} \end{array} \right\} \tag{16}$$

x_i denotes an allocated metric that may be throughput, throughput times hops or response time, and n presents the total number of flows. An index value of F equal to

one is perfectly fair. From the curve, we have founded a fairness value equal to 0.8, and this means that the useable bandwidth is better enhanced and well distributed between nodes. On the other hand, the DAMT algorithm is more efficient for data transmission and this is great result forcefulness and its adaptability to rapid network condition change.

5 Conclusions

Finding a solution for the problem of the loss of communication between sensor nodes, we have proposed an algorithm which increases the lifetime duration of these nodes in order to not disconnect them from the network and do not lose their collected data. We have shown that depending on the traffic condition, the number of slots in a data frame can be modified due to a dynamic adjustment of the CAP and CFP periods. To minimize the packet transmission delay, we have set different priority levels to minimize the over-listening of the channel which often causes a waste of energy. However, with the dynamic adjustment of the duration of the superframe, it is not always possible to connect all sensors all the time to the network. To enhance our proposed algorithm, we think to add in our future work, mobile nodes in the topology that move regularly near the unconnected devices to collect their data.

References

1. Bouazzi, I., Bhar, J., & Atri, M. (2018). A dynamic adaptation mechanism for traffic conditions in wireless sensor network. *Wireless Personal Communications, 101*(4), 1967–1982.
2. Wang, J., Ren, X., Chen, F., et al. (2015). On MAC optimization for large-scale wireless sensor network. *Wireless Networks, 22,* 1877. https://doi.org/10.1007/s11276-015-1073-2.
3. Patel, A., & Upadhyay, R. (2013). Performance analysis of Slotted CSMA/CA MAC protocol under different parameters for static IEEE802.15.4 WSN, Tuscany Region under the Pegaso project, 04-08.
4. Xia, F., Li, J., Hao, R., Kong, X., & Gao, R. (2013). Service differentiated and adaptive CSMA/CA over IEEE 802.15.4 for Cyber-physical Systems. *Scientific World Journal*. PMC3821941.
5. Bouazzi, I., et al. (2017). Priority-based queuing and transmission rate management using a fuzzy logic controller in WSNs. *ICT Express*. https://doi.org/10.1016/j.icte.2017.02.001.
6. Kim, J., & Lee, J. W. (2017). *Telecommunications Systems, 64,* 293. https://doi.org/10.1007/s11235-016-0176-0.
7. Mainetti, L., Vincenzo, M., & Patrono, L. (2015). HEC-MAC: A hybrid energy-aware cross-layer MAC Protocol for wireless sensor networks. *International Journal of Distributed Sensor Networks.* Article ID 536794, 12 p.
8. Subramanian, A. K., & Paramasivam, I. (2017). PRIN: A priority-based energy efficient MAC protocol for wireless sensor networks varying the sample inter-arrival time. *Wireless Personal Communications, 92,* 863. https://doi.org/10.1007/s11277-016-3581-5.
9. Sherly, L., Annabel, P., & Murugan, K. (2015). Energy-efficient quorum-based MAC protocol for wireless sensor networks. *ETRI Journal, 37*(3).

10. Rajesh, B., Shyamala, D., & George, A. (2015). Hybrid CSMA/CA-TDMA Based MAC protocols for wireless sensor networks. *ARPN Journal of Engineering and Applied Sciences, 10*(4).
11. Chen, Z., Liu, A., Li, Z., et al. (2016). Distributed duty cycle control for delay improvement in wireless sensor networks. *Peer-to-Peer Networking and Applications, 10,* 559–578. https://doi.org/10.1007/s12083-016-0501-0.
12. Zhang, H., Cheng, P., Shi, L., & Chen, J. (2015). Optimal denial-of-service attack scheduling with energy constraint. *IEEE Transactions on Automatic Control, 60*(11), 3023–3028.
13. He, S., Chen, J., Jiang, F., et al. (2013). Energy provisioning in wireless rechargeable sensor networks. *IEEE Transactions on Mobile Computing, 12*(10), 1931–1942.
14. Deepak, K. S., & Babu, A. V. (2016). Energy consumption analysis of modulation schemes in IEEE 802.15.6-based wireless body area networks. *EURASIP Journal on Wireless Communications and Networking*. https://doi.org/10.1186/s13638-016-0682-5.
15. Oh, H., & Azad, M. A. K. (2016). A big slot scheduling algorithm for the reliable delivery of real-time data packets in wireless sensor networks. In *Wireless communications, networking and applications. Lecture notes in electrical engineering* (Vol. 348). https://doi.org/10.1007/978-81-322-2580-5_2.
16. Guerroumi, M., Pathan, A. S. K., Derhab, A., et al. (2016). MMSMAC: A multi-mode medium access control protocol for wireless sensor networks with latency and energy-awareness. *Wireless Personal Communications* . https://doi.org/10.1007/s11277-016-3726-6.
17. Wang, H., Chi, X., & Zhao, L. (2016). Energy adaptive MAC protocol for IEEE 802.15.7 with energy harvesting. *Optoelectronics Letters, 12,* 370. https://doi.org/10.1007/s11801-016-6163-6.
18. Sreejith, V., Suriyadeepan, R., Anupama, K. R., & Lucy, J. (2016). DS-MMAC: Dynamic schedule based MAC for mobile wireless sensor network. https://doi.org/10.1145/2851613.2851999.

Depth Estimation Using Blob Detection for Stereo Vision Images

R. Arokia Priya Charles and Anupama V. Patil

Abstract Object detection is one of the most important research areas in image processing. Here, we propose to employ dual camera system to precisely find the distance, the height, the width of object that are applicable in areas like autonomous system navigation, vehicle navigation, etc. One of the most popular approach is the passive approach; however, its performance is usually limited by complicated scenes and texture-fewer areas. Therefore, actual and resourceful estimation algorithms of stereo depth have been a long-time problem. To solve those problems, an enhanced methodology that embraces three steps are proposed. Foremost, the images are segmented using thresholding technique, and a blob detection technique is used to find the objects in the scene. Then, the center of gravity is found for the objects. It helps to avoid the rectification process. Finally, the distance of the objects is found using the disparity values of the objects. Experiment results are promising.

Keywords Stereo vision · Depth map · Blob detection · Center of gravity

1 Introduction

Humans have the capability to coarsely guess the size with its distance of an object due to their double vision. The complications of scene analysis are felt due to innate depth ambiguities in a single camera 2-D image. Huge research works have been done to estimate utilizes only a single vision sensor. Of late researchers have started using multi-vision cameras for such applications. Using multi-view cameras increases the dependency of hardware, and the processing time in relative manner is also more. Hence, multi view systems become costly. On a compromise, stereo vision solves the issues related to monocular vision and the multi-view system. An instance is the

R. A. P. Charles (✉)
Vishwakarma Institute of Information Technology, Pune 411044, India
e-mail: prinnu@yahoo.com

A. V. Patil
Dr. D. Y. Patil Institute of Engineering Management & Research, Akurdi, Pune, India
e-mail: anupamav4@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_22

work of A-Lin et al. where they employ dual camera to calculate driving distance of a vehicle so that it is safe. The distance was measured centered on the inequality of the two frames captured by the dual camera. Later, the effort by Baek et al. offered an progress in the scheming of the inequality such that a additional precise entity distance can be found [1]. Over here in this paper they had revealed that the inequality of the entity in a bigger assessment area can be gained by means of their method. Individuals effortlessly see depth in 2-D images and stereo vision affords a new straight way of inferring the 3-D range and is also profoundly used by the human visual system. Depth input resultant from dual camera has been used beforehand for assortment of uses, such as passive navigation, cartography, surveillance [2]. The increasing use of digital cameras in everyday life poses fresh challenges and interesting aspects that need to be addressed. Improving three-dimensional depth from pictures is an elementary problem in computer vision and has vital uses in audiovisual editing, robotics, scene understanding and 3-D reconstruction. Nevertheless, dual camera difficulty has been considered for a lengthy time; superior solid depth input is still focal debate issue of dual camera quest area, because of some intrinsic difficulties, like noises, regions without any texture, and occlusions. The idea of stereo vision is after all simple by taking pictures of the same scene from slightly different positions, and this difference in their pixel position is called disparity. Disparities between the images can be used to approximate the depth of objects in the scene. A complete map of the depths of all objects in the scene is generated by matching each region of one image with the second and determining the distance these matched regions are from one another. A brief summary on dual camera matching can be seen in [3]. Such usual matching algorithms are classified into local and global methods. Local approaches utilizes the color or intensity values within a certain window to decide the inequality for each pixel even if they need huge extent of calculation, as of their search model and search strategy during pixel matching, like three-step search strategy, four-step search strategy, and diamond search [4]. To decrease the calculation complication of pixel matching, a new examination model of object tracking using blob detection is proposed. This paper is ordered as below. Section 2 gives implementation methods in order to find disparity map. Section 3 describes computing time for CPU and GPU details of our strategy. The experiment results are shown in Sect. 4.

2 Disparity Flow

2.1 Flow With and Without Background

Many methods based on similarity-based techniques for matching the objects through region-based segmentation, and the clustering algorithms are done as per the survey. Here, the focus is on the discontinuity-based match process that is the edge detection technique. In order to detect the edge, we propose to use the blob detection. Identification of real-life objects belonging to many classes in images/videos is called

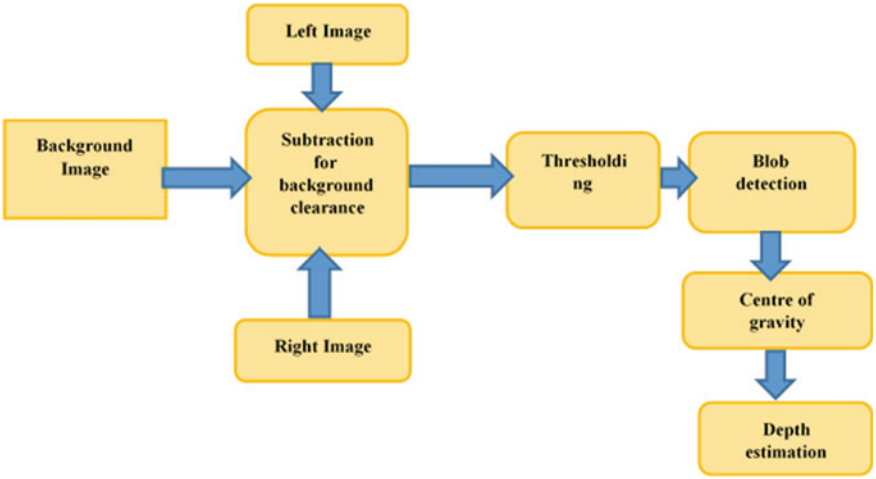


Fig. 1 Process flow

object detection. It is the process which detects the interested objects and discards the unwanted background from the image. Extraction of object is one of the main process/steps in object detection. Objects of interest can be extracted from such an image using connected component or the blob extraction method. We have done the process with and without the background. Region matching algorithms consumed ample amount of time for execution of the task, and so we have used the segmentation process that comes under the region based for the separation of background from the foreground (Fig. 1).

2.2 Algorithm

1. Load input 2-D image from database for both left and right images.
2. Convert input image to grayscale image.
3. The threshold image is applied with the connected components algorithm, i.e.,
4. BLOB is implemented in a binary image, connected non-black pixels are isolated by individual blobs, and array of information is provided.
5. The center of gravity is calculated in order to identify the object, and based on that the disparity is calculated, and the depth is estimated using traditional method. Algorithm with background requires following extra steps. As second step, load the background image separately (the image/scene without any object/first frame) then subtract the background image from the left as well as the right image the background. The output of this is a left and right image without a background, or it is a white background. By subtracting the extracted binary foregrounds from the current frame backgrounds are found [1]. All other steps are the same.

Good amount of information is required for conversion from a color image to a grayscale image. The number of bits the device could support and determine the quality of the color image.

2.3 Thresholding

Thresholding is the simplest method of segmenting image. Binary images can be generated/created from a grayscale image by using thresholding. Every pixel with an image density is replaced either by a black pixel or a white pixel depending on its intensity value (lower—black, higher—white) when compared to a predefined/constant pixel value. The IDE we used is NetBeans along with JAVA coding to make use of multi-threading. Histogram shape-based methods that take into the account the max top value, max bottom value, and curvatures of the smoothed histogram are analyzed. Based on which thresholding is done in order to segment the images as background and foreground. Simultaneous/concurrent execution (multi-threading) of two or more small programs (threads) of a program is a Java feature, which maximizes the CPU utilization.

If $I\,x, y < T$, then $I\,x, y = 0$, else $I\,x, y = 1$

An image subtraction algorithm is used in this process to remove the background of the image in order to properly identify the object of interest. The foremost reason for using this algorithm is its simplicity and can be easily deployed in limited real-time processing. It turned out to work acceptably well. Before looking for the optimal subtraction, we considered a basic subtraction of the images, considering the camera to be still [5]. In case of still camera, the background image is expected to be the same always, especially in applications like security cameras. So we need to perform some basic operations of subtraction as the images are considered to be the still images. In the first go, the images are clicked for background and save in the database, and this case is assumed for no change of background throughout. Any change in the picture that is taken will not give a background subtraction to be zero if there is a change of the input image. Here, we have assumed the pictures are taken in frequent interval of times, and the subtraction happens accordingly. The background frame is called the reference frame. The reference frames are always the same. Any change in input for example a person seen in camera in the left image will be subtracted with the background image stored for left background, and the right image is subtracted from the right background. Here, we will have two background images stored in the memory/database because the camera is a stereo vision camera. Emphasis is on choosing the best frame as reference. Degradation of the border to "poorest sighted border" is undesirable for hence will definitely reduce the signal to noise ratio. Later, for matching process, the reference frame is convolved with the current frame. This is done in order to get an optimal result (Fig. 2).

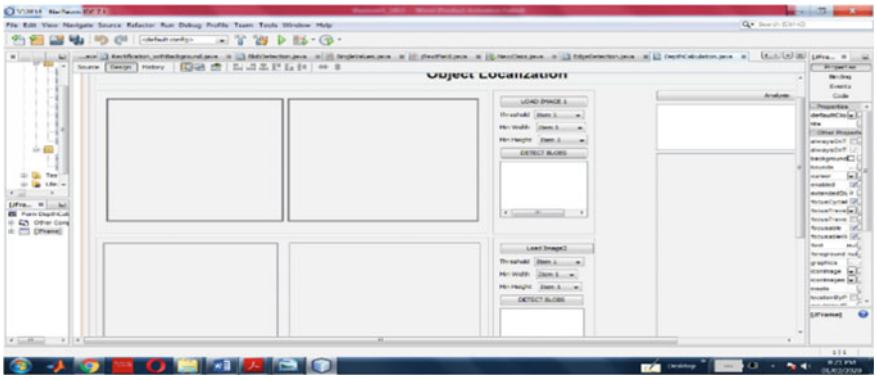


Fig. 2 Thresholding option for segmentation

2.4 Algorithm for BLOB Detection

Blob classification is to classify the extracted objects in both the scenes and to find the object of our interest. Let $b1$, $b2$, $b3$, …, bn be the number of objects with H, W be the height and width, respectively, (Figs. 3 and 4).

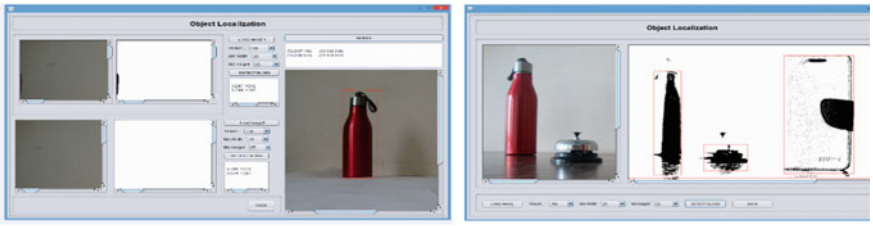


Fig. 3 Blob detection for single object, for multiple objects

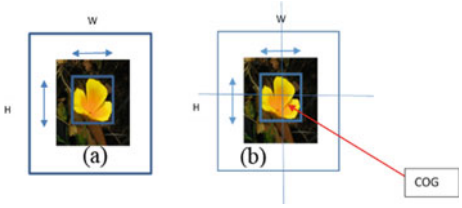


Fig. 4 **a** Input image with H and W. **b** Center of gravity

1. Read the black and white pixels in image,
   For q=0 to ht, For p =0 to wth, Find pixels that are black, Read till next pixel is found, Mark coordinates, End for, End for.
2. Read all values, For value is 1 to 1
3. Mark a rectangle
4. Draw a rectangle
5. End for loop

So *CXi* and *CYi* be the center of gravity points for respective objects.

$$\text{COG} = \sqrt{(x2\text{mid} - x1\text{mid})^2 + (y2\text{mid} - y1\text{mid})^2} \tag{1}$$

$$CXi = Wi/2 \tag{2}$$

$$CYi = Hi/2 \tag{3}$$

2.5 Image Moments

Blob detection algorithm worked well by drawing a rectangle around every object and by finding the center of gravity. But if the blob is outlined with always a rectangle or a square, then the center of gravity method works well. But because it is stereo image there is a difference in the size of objects that are outlined which affects the center of gravity value as the disparity value is not right because of the size in rectangle. If the distance of the object is increases, then this error also increases. So center of mass, for moment is used, and the centroid is given by [6]

$$\overline{x} = M10/M00 \tag{4}$$

$$\overline{y} = M01/M00 \tag{5}$$

3 Computing Time

This method takes all pixels one by one for and lands up in a larger time-consuming operation. But the genuine cost of the calculations is much lesser than it appears to be at primary scan [7]. Utmost of the calculating time is taken by the calculation of the matrix MXN size will take MXN time of calculation (where M is row, and N is the column of image matrix). This process costs the CPU, as it has to do twice

one for left image and one for right image. The whole process gets doubled, i.e., MXNX2. The computing time needed for the remainder/balance of the process is almost the same. By using the proposed method, we can develop a 1000 × 750 frame in about 170 ms with a 2.5 GHz PC. Further improvement and refinement are done by using the concept of multi-threading in JAVA. The process time was reduced to almost half while multi-threading was used because the threads helped in doing the subtraction simultaneously. It gave a time consumption of (MXNX2)/2 = MXN. In order to further improve the consumption time, the whole process was executed in multi-core processors using GPU. The GPU configuration is NVIDIA Quadro. K1100M with Warp size = 32, Max Threads = 2048/SM, Number of cores: 384, i.e., 192/SM, Number of SMs = 2, Threads/block = 1024.

4 Result and Discussion

To test the efficiency of our approach, the outcome of research shows that our technique can realize detection of moving objects effectively in the indoor environment with a lighting effect change. Real-time object detection using the video inputs was done. USB camera Logitech C270 HD, with specifications of Max Resolution: 720p/30fps, focus type: fixed focus, lens technology: standard, built-in mic: mono, FoV: 60°, is used. The video input from both cameras is taken. There was an image frame lag between both the camera inputs as the code was executed sequentially, and hence, the frames with objects were not matching. It incurred as huge difference in the disparity, and at times the frame itself was $(n + 1)$th frame. Hence, multi-threading concept was used for image acquisition, and this gave a great result in terms of time consumption and the frame matching. Figure 5a, b is the images used as a background image, and Fig. 5c, d is input images of left and right camera. For the blob detection with background subtraction, this background is considered and the subtraction is done as per the flow explained earlier in this paper. Figure 6a shows the input kept at a distance of 30 cm from the camera, and the cameras were 50 mm apart, (c), (d) shows the segmented output from both the cameras as per the process explained earlier, (e), (f) gives the centroid of the objects, and (g), (h) shows the blobs that are detected from which the depth is calculated. As the object distance

Fig. 5 **a** Left background. **b** Left single object. **c** Right background. **d** Right single object

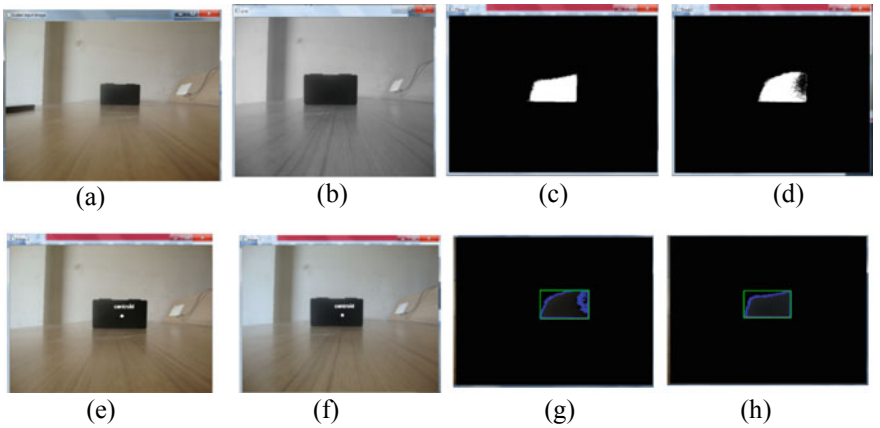


Fig. 6 **a** Input with actual distance = 30 cm. **b** Gray scale image. **c**, **d** Segmented images left and right. **e**, **f** Single object with centroid left and right images. **g** BLOB left. **h** BLOB right

increases, since the moments do not take care of the scale, translation, and rotation invariant property, the centroid value calculated was not accurate. The accuracy kept on decreasing with the change in object distance. More the distance less the accuracy (Table 1).

Depth is computed using the very classical triangulation method in which f is the focal length of the camera system; B is the base line ($x2$–$x1$) is the disparity;

Table 1 Depth using centroid (baseline 50 and 75 mm) and its execution time

Actual depth (cm)	Baseline 50 mm		Baseline 75 mm	
	Estimated depth (cm)	Error in cm	Estimated depth (cm)	Error (cm)
43	46.4876	3.4876	49.48	6.4876
53	53.15249	0.15249	57	4
63	62.33039	0.66961	66.32	3.32
73	72.69385	0.30615	75.85	2.85
83	83.78274	0.78274	86.27	3.274
93	93.26758	0.26758	100.3	7.328
103	102.983	0.017	110	7

Table 2 Execution time of CPU Vs GPU

Images size	CPU	GPU
3264 × 2448	In minutes	In µs
4288 × 2848	In minutes	In µs
1000 × 750	In secs	In µs
1000 × 664	In secs	In µs

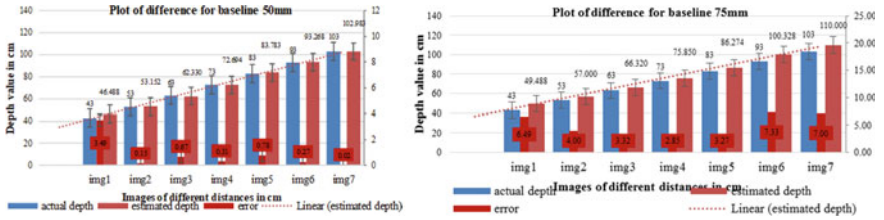


Fig. 7 Plot for baseline of camera as 50 mm with error and 75 mm with error

normally by looking at the specifications of a camera module, we obtain the focal length in mm or cm but in order to compute depth in cm or mm; we have to convert this focal length from mm or cm to pixels.

$$f \text{ in pixels} = f(\text{mm or cm}) * \text{image width(pixels)}/\text{CCD size(sensor width)} \quad (6)$$

We experimented with keeping the cameras apart at 50 mm and other at 75 mm. From the above two plots, we can see that the baseline 50 mm gives the error value to be 1–3 cm, and the baseline with 75 mm gives 2–7 cm. As we increase the distance between the cameras the depth error also increases. As the distance between the eyes is 30–50 mm, this algorithm would suit for robotic kind of applications (Fig. 7).

5 Conclusion

We determine the calculation of moment invariants for the distances from 30 to 110 cm in 3 cm increments in Lenovo ThinkPad W541 (20EF000NUS) Laptop: i7-4810MQ (up to 3.8 GHz), the moment invariants change as distance between the cameras increase, because images size decreases when the object distance increases, and hence, moments work much better in shape matching as the shape of the object won't be that significant. This does not work for more number of blobs, especially if the objects overlap. From the trial readings, the choice of image spatial resolution is found; the distance between the cameras is vital to preserve invariant features. To lessen the computation time of moment invariants, the image spatial resolution must be lower in case of CPU. However, the resolution can be too high in case of GPU, since the calculation will extraordinarily rise as the resolution escalates. Hence, an appropriate resolution is to be chosen to get an uncompromising balance between calculation and resolution in actual use.

References

1. Jadhav, L. H., & Momin, B. F. (2016). Detection and identification of unattended/removed objects in video surveillance. In *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology*, Bangalore, India (pp. 1770–1773).
2. Medioni, G., & Nevatia, R. (1985). Segment-based stereo matching. *Computer Vision, Graphics, and Image Processing, 31*(1), 2–18. https://doi.org/10.1016/s0734-189x(85)80073.
3. Hamzah, R. A., & Ibrahim, H. (2016). Literature survey on stereo vision disparity map algorithms. *Journal of Sensors, 2016,* 1–23. (Hindawi Publishing Corporation).
4. Chau, L.-P., & Jing, X. (2003). Efficient three-step search algorithm for block motion estimation in video coding. In: *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)* (pp. 421–424). https://doi.org/10.1109/icassp.2003.1199501.
5. Desa, S. M., & Salih, Q. A. (2004). Image subtraction for real time moving object extraction. In *International Conference on Computer Graphics, Imaging and Visualization*, Penang, Malaysia. https://doi.org/10.1109/cgiv.2004.1323958.
6. Rocha, L., Velho, L., & Carvalho, P. C. P. (n.d.). Image moments-based structuring and tracking of objects. In *Proceedings. XV Brazilian Symposium on Computer Graphics and Image Processing*. https://doi.org/10.1109/sibgra.2002.1167130.
7. Alard, C., & Lupton, R. H. (1998). A method for optimal Image subtraction. *The Astrophysical Journal, 503*(1), 325–331. https://doi.org/10.1086/305984.

Augmented Reality in Visual Learning

Suma Karanth and Rajashekara Murthy S.

Abstract Augmented reality (AR) is one of the upcoming technologies in the area of information technology. It is gaining wide momentum and is being deployed in various fields of advertising, marketing, manufacturing, health care, tourism, etc. With the increase in the usage of smart phones, Internet and cloud connectivity, mobile technology and augmented reality can be integrated to provide a richer learning experience in the field of higher education and research. Complex concepts, practical hands-on and laboratory experiments can be captured within customized AR application to enhance learning and provide in-depth understanding to students. This paper elaborates customized software design and development for mobile-based AR application to capture details specific to Indian medicinal plants. This application deployed on smart phones overlays digital information when a smart phone camera points to specific botanical plants. This application relies on Vuforia's Image Target recognition that is a type of marker-based algorithm.

Keywords Augmented reality (AR) · Software development kit (SDK) · User interface (UI) · Machine learning (ML) · Image Target · Vuforia

1 Introduction

Augmented reality (AR) is one of the upcoming technologies in the area of information technology. It is gaining momentum and is being experimented in various fields of marketing, education, manufacturing, health care, automotives, entertainment, tourism and much more. It is one of the key technologies mentioned in Industry 4.0 standard. With the increase in the usage of smart phones across length and breadth of India, Internet, mobile technology and augmented reality-based applications can

S. Karanth (✉) · R. Murthy S.
R. V. College of Engineering, Bangalore, India
e-mail: karanthsumas.sit18@rvce.edu.in

R. Murthy S.
e-mail: rajashekaramurthys@rvce.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_24

go hand in hand and provide a richer user experience in the fields of education and research. Complex concepts and practical hands-on can be captured in the form of AR-based application to provide in-depth understanding to users who could be students, researchers and hobbyists.

Educational fields, especially science, technology, engineering and mathematics, i.e., STEM-based fields, can benefit a lot from this technology if AR software is developed in a well-designed manner and applied in relevant situations [1]. Though Internet provides huge number of resources for learning, it can be overwhelming for students and can cause deviation from subject matter; in such circumstances, customized application specific to classroom or laboratory specific training can be built and deployed which can reduce the learning curve and avoid distraction. It can help users understand complex concepts or practical areas in a better way through use of 3D models, images, annotations, videos specific to the subject under study and provide on the field learning experience.

Identifying plants, their botanical names, their family classification, order, native origin form a part of botanical studies. Cultivation of Indian medicinal plants is gaining momentum. An independent "National Medicinal Plants Board," i.e., NMPB has been set up by Government of India to coordinate matters related to plants [2]. By utilizing AR technology to identify and project more information about various plant species, institutes such as NMPB can be benefitted.

This paper presents a prototype application, which caters to the study requirements of botany and considers sample medicinal plants commonly used in Indian household. This application overlays digital information when a smart phone camera points to specific flash cards of these botanical plants such as Tulsi, Aloe Vera, Amla, Brahmi, etc. These digital objects include textual information such as botanical names, family, order; video lectures specific to these plants, etc. Details about these plants are prepared in advance and are predefined in storage of application as this is based on marker-based algorithm for AR invocation. It requires on the field analysis by capturing snapshots of good quality of plants that need to be studied.

Initial section of the paper covers a basic understanding of AR technology, hardware and software requirements and available software platforms to implement such applications. It also considers criteria used to select software platform for development and deployment.

Next sections consider the problem statement used for study, software requirements, design and development followed to implement the prototype application. It covers implementation details at a high level as well as screenshots of developed application.

2 AR Technology

Augmented reality is one of the modern technologies in information technology field, which seamlessly blends the digital and physical or real world. AR implies supplementing real-world surroundings with digital objects that can be viewed via various devices such as smart phones, head-mounted displays (HMD), web browsers, etc.

In the visual perception domain, AR helps us to augment 3D virtual content to the real worldview when viewed through a camera [3]. The augmentation of virtual content happens in perspective and in real time giving an illusion that the virtual content is actually in the real world [3].

The main factor in AR design includes tracking and object invocation. There are different methods to track surface on which digital objects can be superimposed. These include methods such as marker-based algorithms, natural feature-based or marker-less algorithms, model-based algorithms.

In marker-based algorithm, 2D patterns having unique patterns and shapes are placed on the real objects, where the augmentations are superimposed. These involve placing flash cards with predefined images or posters near the object of study. Machine learning (ML) methods can be used for image classification and labeling, and these in turn can act as a trigger to activate AR. This is a form of marker-less algorithm. Model-based algorithms are based on 3D model object recognition and invocation.

Following sub-sections cover the building blocks or components of an AR application considering marker-based approach to tracking and object invocation. Figure 1 depicts AR workflow at a high level.

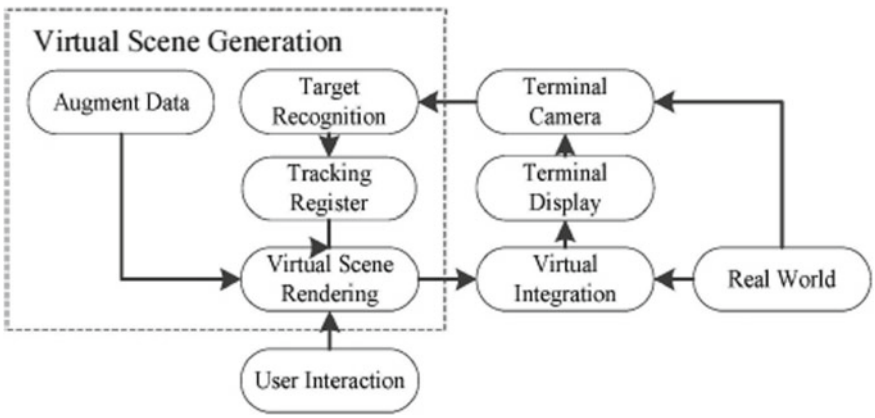


Fig. 1 Steps in AR recognition and rendering [4]

2.1 Camera

Smart phone camera is used to focus on image for target recognition. Predefined image targets are placed in the camera view in the form of flash cards or printouts.

2.2 AR Software Development Kit (SDK)

It connects various layers of control flow from image recognition to actual digital image superimposition. Vuforia as well as Google AR Core provides Software Development Kit (SDK) to design and develop mobile-based AR application. AR application based on Vuforia/Google AR SDK takes the display screen of mobile devices as the hub to connect the AR world.

2.3 Predefined Database

It provides database component to hold predefined image targets. These images are rated based on quality of contents. Higher the rating, more acceptable is the image for target recognition.

2.4 Assets

Assets imply 3D models, 2D images or UI elements such as texts, virtual buttons, graphical widgets, videos that can be superimposed in camera view of user. These assets provide relevant information specific to the Image Target or object under study. Designing these assets requires understanding of design software such as Maya, Blender. Such assets are also available for download in online repositories. These can be paid or unpaid assets.

3 Software Platform Selection

As part of the requirement analysis of the prototype project, Google ARCore and Vuforia SDK were compared before proceeding with the design and implementation. This required some initial analysis as well as few experiments using both SDK platforms. These are some of the observations with respect to usage of both the SDKs.

Google ARCore is based on concepts of motion tracking, environmental understanding and light estimation. Augmented images is a feature that allows you to build AR apps that can respond to specific 2D images such as product packaging or movie posters [5]. Users can trigger AR experiences when they point their phone's camera at specific images—for instance, they could point their phone's camera at a movie poster and have a character pop out and enact a scene [5]. Considering Google's product releases in the area of ML and cloud platform services, intertwining AR along with these latest technologies can be helpful in developing advanced AR applications.

However, ARCore and Sceneform plug-in require latest version of Android; phones with Android version 7.1.1 level and above are useful. Also, for some of the feature implementation, only latest and certified phones can be used for AR application deployment. These phones are listed in following Google ARCore website [6].

Vuforia SDK supports all AR features and is focused in the area of AR development from quite some time. Vuforia along with Unity 3D software can be used to develop applications across multiple platforms. Phones with Android version 5.1.1 and above can be used to deploy AR applications. The limitation lies in licensing of application, when developed for commercial use.

From educational perspective (i.e., non-commercial purpose), Vuforia SDK integrated in Unity 2018 provides relevant features for design and development of this project. The developer community support for Unity–Vuforia platform is widely available to address common issues faced during development of features.

4 Related Work

This section provides step-by-step methods used to design and implement the mobile-based AR application.

4.1 Scope

Currently, users study botanical information theoretically, and practical information is provided orally through experts. There is a scope to digitalize this approach and add more to user's experience by providing an AR application.

4.2 Methodology

Developing an AR application requires hardware as well as software support. "Predefined data" is one of the main aspects to develop this application. This project is implemented using iterative lifecycle model of Software Development Life Cycle (SDLC).

- Hardware in the form of paper or flash card cutting with good-quality images pasted over each of the clipping is assigned to the plants that are under study.
- Vuforia platform is used to capture predefined image targets in database. Vuforia SDK incorporated within Unity 3D is used to develop and compile mobile AR software.
- Field study is required to identify plants that need to be considered. Sample images and theoretical data need to be collected from actual site and concerned staff members.

4.3 Requirements

Following requirements are listed at the start of the project. These include functional as well as performance requirements.

- Application has a meaningful launcher icon and splash screen associated with it; camera view is enabled when user selects to study a subject
- Application allows user to focus camera on predefined tags or image targets
- On Image Target identification, associated 2D/3D assets and/or texts are augmented within camera view; user can focus across multiple objects, one at a time
- When user moves camera view out of scope, augmented objects are removed
- When user closes the application, switch off camera with minimal delay
- This application takes less time for startup once user presses associated launcher icon
- Target object should be recognized with minimal delay (~30 s)
- On recognition, processing of object and display of augmented data should be within feasible time limit (~2–15 s).

Following tables, Tables 1 and 2 capture the software and hardware requirements specific to this project.

Table 1 Hardware requirements

Devices	Desktop web camera or Android smart phone, version 7.1+
Target images (tags)	Hard copy of images for target recognition, tied to or placed near specific plants for study

Table 2 Software requirements

Operating system	Windows 7
Integrated development environment (IDE)	Unity 2018.3.12.f1
Software development kit (SDK)	Vuforia (in-built in unity version 2018.3.12f1 Personal) Android SDK and tools, Java development kit (JDK)
Repository	Git and GitHub

4.4 Design

The components of AR application at a high level include camera interfacing, image processing, comparison of processed image with image targets in precompiled database and rendering virtual objects at tracked position. These virtual objects can be observed in the camera view.

In Unity 3D, all pages with respect to mobile application correspond to Scene. Every Scene consists of objects called as GameObject. AR Camera is one of the GameObject components, which is imported from Vuforia package. Every Image Target is a Vuforia component that can be associated with predefined image compiled in Vuforia database at their developer portal [7].

Classes used in the application are derived from class "MonoBehaviour" as shown in Fig. 2. Every Unity script derives from this base class named "MonoBehaviour". Start() and Update() are two methods which are overridden to achieve desired action such as activating or deactivating GameObject instances.

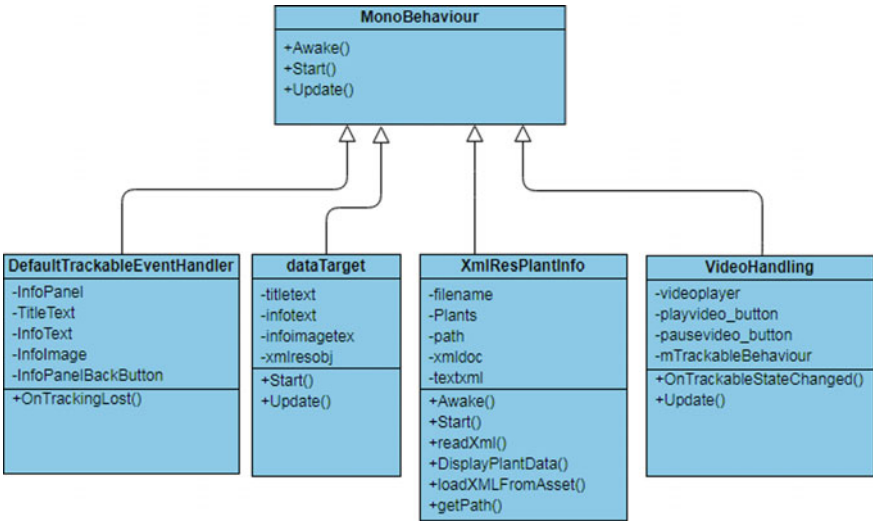


Fig. 2 Classes in Unity 3D scripting

```
<?xml version="1.0" encoding="utf-8"?>

<Plants>
    <Plant>
        <name>tulsi</name>
        <title>Information about Tulsi</title>
        <info>
        Common name: Tulsi, Holy Basil
        Scientific name: Ocimum tenuiflorum
        Family: Lamiaceae
        Order: Lamiales
        Native: Indian subcontinent, Southeast Asian tropics
        </info>
        <infoimage>Tulsi</infoimage>
    </Plant>
```

Fig. 3 XML file for plants

To make the design maintainable and accommodate more plants in the long run, XML file is maintained which consists of plant specific information in each node and associated attribute. This XML file is read at the start of the associated script, and data is updated in respective UI elements of Canvas based on specific plant recognition.

4.5 Implementation

In Unity Editor, "Inspector Window" for every Scene captures the properties and attributes of GameObject instances. Spatial relationships need to be explicitly set between GameObjects. We attempted to remove this step and generalize spatial relationship using XML file and associated parser for all plants, which is extensible.

HomeMenu: This Scene provides options for user selection and allows transition to initiate graphics mode or video mode.

GraphicsModeScene/VideoModeScene: At the start of the Scene, XML file is read as a Text Asset. Details of all plants are loaded into member variables. This Scene consists of AR Camera instance. On activation, based on availability of Image Target in camera focus, associated Canvas UI elements are updated with plant details. Associated video lecture can be played or paused. Figure 3 illustrates example of data captured in XML file specific to a plant.

4.6 Outcome and Deployment

This section captures some snapshots of prototype application. Unity provides facility to build *.apk file for Android smart phone. Player settings need to be updated by

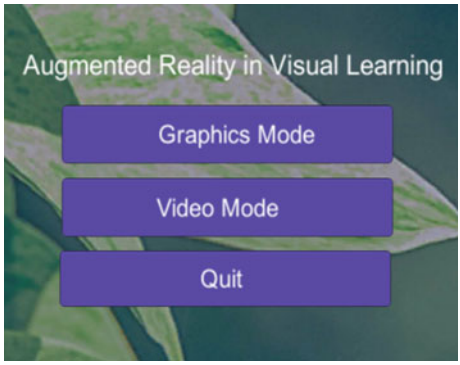


Fig. 4 Main menu screen

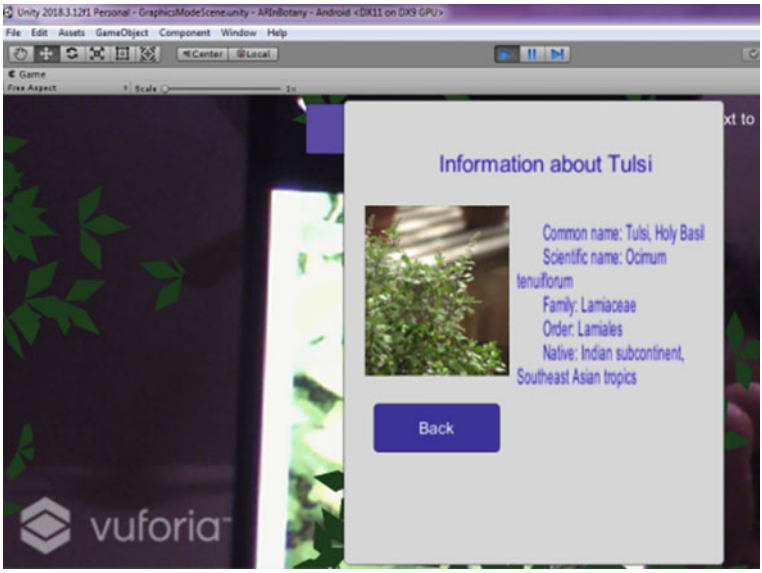


Fig. 5 Canvas UI with details about "Tulsi" plant

selecting minimum API level for Android phone. Associated Android SDK and Build Tools as well as Java SDK need to be mentioned in Unity Editor (Figs. 4 and 5).

5 Enhancements

With respect to introduced prototype application, third-party plug-ins such as Lean can be used to rotate, scale and transform augmented objects. Video can be played

by providing URL. Marker-based application can be used in laboratories or garden areas when markers or hard copy images of predefined image targets are placed in the camera view of the application.

In places where markers cannot be placed and insert, marker-less-based algorithm can be used by incorporating machine learning methods. Tensorflow provides facility to build models based on images of various types and in various dimensions as well as attributes. Based on accuracy of image labeling, associated assets such as text transcripts can be placed using AR platform. This implies combining machine learning and AR technologies.

6 Conclusion

AR technology is being applied in interesting ways in many domains and no longer used only from gaming perspective. Latest innovations in the area of deep learning are opening up new possibilities of intertwining AR and ML. Its engaging nature has been considered for development of various advertisements such as "IKYA" furniture, real estate ads, etc.

The prototype application discussed in this paper can be used to enhance knowledge of students, researchers, hobbyists to showcase complex objects and their internal workings such as plant structure, botanical name, family, order, native origin, etc. This project is based on marker-based recognition for invocation of augmented objects. Lot of research and development work is going on in the area of marker-less technology as well as dynamic invocation of AR models. Google has released further revisions of ARCore/Sceneform plug-in; Vuforia is continuing with enhanced SDK features to make AR more beautiful and engaging in various spheres and markets.

Such innovative methods can be applied in various fields where complex concepts and practical knowledge need to be grasped thoroughly and clearly. It can provide precise information on the field and avoid overload of information. Overall it is a powerful technology which needs to be adopted in a balanced manner, such that it does not intimidate real-life experiences.

References

1. Ang, I. J. X., & Lim, K. H. (2019). Enhancing STEM education using augmented reality and machine learning. In *2019 7th International Conference on Smart Computing & Communications (ICSCC)*. 08843619 IEEE.
2. https://www.nmpb.nic.in/.
3. https://www.cdac.in/index.aspx?id=pe_ar_blr_augmented_reality.
4. Peng, F., & Zhai, J. (2017). A mobile augmented reality system for exhibition hall based on vuforia. In *2017 2nd International Conference on Image, Vision and Computing*. 07984714 IEEE.
5. https://developers.google.com/ar/discover#how_does_arcore_work.

6. https://developers.google.com/ar/discover/supported-devices.
7. https://developer.vuforia.com/vui/auth/login.
8. Sunil, S., & Nair, S. S. K. (2017). An educational augmented reality appto facilitate learning experience. In *2017 International Conference on Computer and Applications (ICCA)*. 08079771 IEEE.

A Study and Analysis of Applications of Classical Computing and Quantum Computing: A Survey

A. V. Navaneeth and M. R. Dileep

Abstract In the world of computation, the invention of Computers has reduced human effort for accomplishing task. Depending on the type, size and nature of the work the computers are manufactured in different magnitudes. As the technology grows, it facilitated to manufacture high performing, high throughput, and high memory Classical computers. Classical computers store any information in the form of bits in the memory device, which can be either 1 or 0 (binary system) and all the processing will be done in the same logic i.e., high voltage/low voltage. Because of this logic the classical computers solve some problems by using more resources. Quantum computer are the next level computers which uses quantum theory principles for solving problems. The quantum computer uses Qubit as basic unit which uses principles of quantum computing such as superposition and entanglement, which gives qubit, the efficiency of exhibiting many logical states at a time. This paper shows a comparative study on classical and quantum computers and illustrates various problems in classical computers and respective solutions with use of quantum computers.

Keywords Qubit · Quantum computing · Superposition · Entanglement · Correlation

1 Introduction

The classical computer uses binary logic for processing any data, and the data are stored in the form of bits. Currently almost all efficient computers built on same basis, and solving all problems with binary logic, however there are some problems where

A. V. Navaneeth (✉) · M. R. Dileep
Department of Master of Computer Applications, Nitte Meenakshi Institute of Technology, Yelahanka, Bengaluru, Karnataka, India
e-mail: avnavaneeth25@gmail.com
URL: https://www.nmit.ac.in/department/mca-faculty.html

M. R. Dileep
e-mail: dileep.kurunimakki@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_25

classical computers consumes enormous amount of time for finding the solution. Such problems are addressed by implementing a new way of computing i.e. quantum computing/computer, which uses the laws of quantum physics, and solves the problem using qubit as basic unit. It is estimated that, Quantum computers are faster in solving the problem than classical computers. This paper demonstrates a comparative study on classical and quantum computers with some unique illustrations.

The primary approach of this paper is to identify the drawbacks of classical computing and analyse it, and transform these drawbacks into the potential ideas. In this regard, the study made over identifying the drawback of the classical computers and it is observed that, certain problems require different dimension of views for solving them efficiently. And as a result, quantum computing is estimated to be a potential tool for such problems. Unique problems are identified and stated in this paper such as chess game, card face, there by an attempt is made, how classical computers face difficulty in solving them, and further, a bird view of quantum computing is demonstrated, by illustrating few problems which can be solved easily by quantum computers, and there by directing the researchers in the area of quantum computing.

1.1 Classical Computers

Classical computers are devices which can solve any problem in terms of digital logic in discrete forms. It uses 0 and 1 as binary logic, and usually inputs will be in the format of numerical, alphabetics, special characters, video and even audio. The processing of data basically involved 2 steps

- Through Arithmetic and logical unit and
- Operation on ALU.

The Organisation or the Architecture of a Classical Computer is shown in Fig. 1.

Classical computers basically receive data in 3 forms.

- File, with a name and type
- Record, with a size
- Field, with a location for storing.

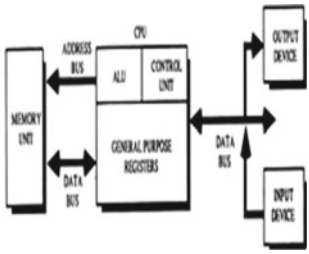


Fig. 1 Processing data in classical computers

Record consists of data about some entity and its properties, operations on the properties, type of the format of data which that entity has etc. The illustration can be the details of a customer in some bank or in any of the originations. If the customer in bank is considered, then customer name, customer id, customer address, phone number, email id etc, are examples of its properties and, this customer details will be in the form of a record. Similar records can be created for number of customers. Classical Computers processes data in 3 steps.

- Preparation of Data
- Programming, and
- Check the validity of the program.

Based on Application, classical computers are grouped in 3 ways.

- In the field of science and technology for processing data
- In business-oriented environments for processing data.
- Versatile Computers.

The computers which are used in science and technology fields majorly used to accomplish task on data which are in different forms, can able to evaluate mathematical equations based on collected data, drawing hypothesis, etc. The computers which are used in this field basically of huge capacities, can have many properties, and can be able to handle data in many forms. Whereas the computers which are used in business applications are capable to gather the data (symmetric or asymmetric) efficiently, processing data statistically, drawing the observation in various forms, and finally taking decisions on them. The illustrations can be Allocation of loan to customer in banking environments, fixing of discounts for products purchase in online marketing etc. The third category is versatile computers, which is combination of both science and technology oriented, and business fields. Most of the modern types of computers are of versatile computers. The Versatile Computers broadly classified in two ways:

- Computer for General purpose
- Computers for Special purpose.

For processing basic scientific applications or mathematical evaluations/calculations, and for performing business operations, general purpose computers are useful. When the requirement of processing, is of particular type then special purpose computers are used, such as

- Controlling of mechanical tools, based on numerical figures.
- Switching operations in electronic devices.
- Guiding the satellites, in finding the path to rotate around the earth etc.

Characteristics of Classical Computers
The following are the characteristics of a Classical Computer.

- Classical computers are efficient at executing the instructions stored at its memory.

- Classical computers are enormously versatile at working environments. They can solve scientific and technical problems efficiently; at the same time, they are good at solving business-oriented problems also.
- As compared to human working efficiency the classical computers have high accuracy rates and the error rates are also adequately less, except for some problems.
- Another most important characteristic of classical computers is that they are very fast, and can solve even large set of problems within short time.

Problem Identification in Classical Computers
Illustration 1:
Problem statement: The problem consists of 4 cards with 2 faces (for each card), face-1 consist of a number (from 2 to 10 or any character among the following: A, K, Q, J) and face-2 is opaque (applicable for all cards for the corresponding face). Assume that face-1 value for all cards are as follows: 2, 3, 4, Q. Index the card which is having face-1 value Q as 1, and all other cards with face-1 value other than Q as 0. Shuffle the cards. Now one of the cards among 4 cards will be having index 1 and all other cards will be having index 0. Find the card with index 1.
Precondition: Load all the index value to a computer memory (classical computer)
Solution: Write an algorithm to find the number with value "1" in a given set of inputs. Assuming that the algorithm is correct,

- The possibility of discovering "1" in the given input at its worst case will be $C(\text{worst})m = m - 1$
- The possibility of discovering "1" in the given input at its best case will be $C(\text{best})m = 1$
- The possibility of discovering "1" in the given input at its average case will be $C(\text{avg})m = m/2$ (approx).

From the above statements it is observed that the algorithms get the desired output with a smaller number of iterations, only when the input value "1" is at first position, for all other valued positions the algorithm is running for multiple times.

The Algorithm failed to find the value "1" at any location with only one try.
Hypothesis: This example proves that the classical computers run for more time by spending its CPU cycles based on algorithm written, for performing a task. It is observed that no co-relation between various values of input, and because of binary nature of the classical computer the output is entirely dependent on the logic, strategy (not considering computer configurations) used in building the algorithm for a given input. The performance can be only increased by implementation of better logic or strategies.
Illustration 2: Computer Chess: Chess game is one of the few arts where composition takes place with performance. Chess is played on an 8×8 alternate black and white square board with pieces and pawns with constrained moves.

Problem Statement: Human players uses various theories, calculations, strategies, logics, on the board by recognizing different patterns, and establishing relationship among the patterns for better game play.

Whereas the way computer plays chess is very different. It identifies all the possible moves and finds the best move among all the identified moves. For doing the task computer chess engines uses tree structures, where all the identified moves are converted into nodes, and each node might be having many sub nodes or children. The tree size increases in proportion with number of moves.
Ex:
1st move (white or black) = 20 outcomes for each side
2nd move = $20 \times 20 = 400$ outcomes
3rd move = $400 \times 20 = 8000$ outcomes
4th move = $8000 \times 20 = 1{,}60{,}000$ outcomes
⋮
If the sequence of moves continues as in the example, at the nth move of the game at some stage, it is estimated that the possible outcomes for the moves may reach around "10^{120}" or even more. It is very difficult to model all "10^{120} outcomes into tree structure and finding (best move) one leaf node in the tree. Even super computers may take years to find the best move from the tree. Because of this, there are many alternate methods are available to calculate the approximate best move.

Based on the context and by observing the scope of this paper, the discussion about computer chess will be restricted here.
Hypothesis: This example proves that, if a problem's order of growth increases exponentially in drastic way, then classical computers takes more time for making computations.

1.2 Quantum Computing

Quantum computing is a new computing method where the computer is designed with use of quantum principles at atomic and subatomic levels. Currently the quantum computer is designed at a small-scale level by some industries. Due to various factors which affect the performance of a quantum computer, the development of quantum computers is restricted for a small size. If quantum computer is developed in large sizes, by considering the suitable environments for its sustainability, then quantum computers are way faster than current super computers. The issue is with the designing of quantum computer i.e., quantum computers use an atom or electron or even a proton as a qubit. To exhibit the quantum features, the qubit has to be maintained in certain atmosphere where all qubits works in correlated condition. Even a minor change in the status of the atmosphere of the quantum system may disturb the qubits, and my lead to breakdown of the system. Currently IBM was successful in designing a quantum computer with 51 qubits. If thousands, or lakhs number of qubits are correlated with each other by any technology for designing quantum computer, then there is no doubt that, quantum computers acquire supremacy. By following the laws of quantum physics, quantum computers exhibit multiple states

at the same time, and establishing a correlation ship among them, there by incredibly enhancing its processing powers.

Quantum Theory

During 1900s Max Planck introduced the Quantum Theory to the German Physical Society. The quantum theory states that energy subsists in individual element, called "quanta", as in matter. After the introduction of quantum theory, number of researches is carried out about the existence of energy, about units, quantity of units, and so on related to quantum theory. It took many years for the scientists to figure out the uses of quantum theory.

The vital rudiments of Quantum Theory

- Energy in a substance will be present in the form of discrete units called quanta, not like in continuous waves.
- Depending on the conditions, the Elementary particles of both energy and matter, might act like either waves or particles.
- The basic substance is considered to be of highly unpredictable nature, due to their inherently random movements.
- The position and momentum of elementary particle cannot be measured accurately at the same time. If position is measured precisely, then there is always a flaw in measuring momentum of the particle. Thus during simultaneous measurement of position and momentum of elementary particle, both the properties stand inversely proportional to each other in term of measurement.

Supplementary Improvements of Quantum Approach

The Copenhagen explanation of quantum approach was proposed by Niels Bohr, which emphasizes that a substance is anything which is calculated, for instance, a particle or a wave, but it can't be supposed to contain particular features, or even exist, till it is measured. In brief, Bohr formulated the impartial certainty does not subsist. This decodes to a theory known as superposition, that prerogatives while it is not known what the status of any entity is, it is fact in every potential states concurrently, as long as it is not checked.

To demonstrate this hypothesis, it is better to use Schrodinger's Cat. In a bulky lead box, place a living cat. The cat is alive in the box at the current stage. Throw a ampul of cyanide in the box and close it. It is not known whether the cat is still living or it has wrecked the cyanide capsule and died. The cat may both alive or dead at this situation, but it is not known that, the cat is dead or alive. If dead and alive are considered as two states, and further if both states are not known, in converse, the cat may be exit in any of the states and a combination of the states, then according to quantum law it is considered to be a superposition of states. The superposition will be lost, only the box is opened and check, in what situation the cat is, and the cat ought to be either dead or alive.

The multiverse or many-worlds theory illustrates the quantum theory. The multiverse theory demonstrates that as soon as latent exists for any entity to be in any situation, the universe of that entity transfigures into a sequence of similar universes equivalent to the number of possible states in which that the object can subsist, with

each universe contains a distinctive single probable status of that entity. The multiverse theory even states that, these multiple universes can able to interact with each other with use of some mechanism, i.e., changes done to the states in one universe can reflect the state of same entity in another universe in some proportion. Late Richard Feynman and Stephen Hawking are the scientists who had articulated a predilection for the multiverse principle. This principle states that, a particle can present in abundant forms raises in philosophical applications for computation.

2 Comparative Analysis

Classical computing works on the principles of Boolean algebra, functioning usually with a 7-mode logic gate standard, though it is possible to exist with only three modes that is, NOT, AND, and COPY. The presentation of data will be exclusively in terms of 2 state logic, that is, either 0 (off/On) or 1 (false/true). The states are represented by binary digits, or bits. At the heart of computer millions of capacitors and transistors will be in one of the states. The state of capacitors or transistors (0 or 1) will be measured in millions of seconds. The speed of classical computers purely relies on, in how much time the computer takes to measure these states. With the technology the classical computer also growing up still. But since classical computers are limited to some constraints in terms of computations, which leads to exploration of new computation methods, and that is how quantum computing started to grow with quantum physics basis.

In quantum computing the qubits can able to be in states 0 or 1 or a super position of 0 and 1. That means a single qubit can able to be in more than 2 states in quantum computer, whereas this is not possible in classical computers. The basic elements for qubits in quantum computers can be an atom, electron, neutrons, ions etc. And these basic elements acquire some state when it is brought under some force by changing it spin or rotations. The most basic force which incurred for changing state is polarization. In quantum computing after deciding the type of element for qubit and the way of changing the states, these qubits should possess some properties like superposition and entanglement.

2.1 Superposition

An electron which is in magnetic field is considered to be an qubit. The spin of the electron may be either in one of the states. If the electron spin is associated with the field, it is called as a spin-up state, and if the electron spin is reverse to the field, then it is called as a spin-down state. Altering the spin of the electron from one state to another is accomplished by means of an energy pulse, such as through a laser or may be by other methods. Assuming, a unit of laser energy is used for changing the spin of an electron, instead of using a unit of energy, instead, if only half unit of laser

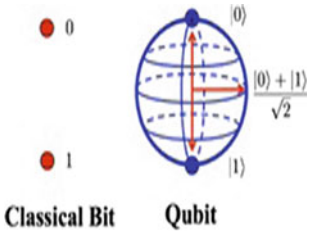


Fig. 2 Classical bit and qubit

power is used, then according to quantum law, the electron enters a superposition of states, where the electron behaves as if it were in both states concurrently. Each qubit which is used can take a superposition of both 0 and 1. Thus, the number of computation that a quantum computer could take is 2 m, where m is the number of qubits used. If a quantum computer comprise of n qubits, then it would have a latent to do 2 m calculations in a particular step. If $m = 500$ then the number of computations would be 2500, is definitely more atoms than the atoms which are there in the known universe. This is called true parallel processing. Today, even though classical computers exhibit parallel nature, but still they are doing one job at a time. There are just two or more classical computers are doing them. But the way the quantum bits interact with each other, would be by quantum entanglement as shown in above Fig. 2.

2.2 Entanglement

Entanglement is a process in which, the element such as photons, atoms, electrons, or qubits that have interact at some summit, maintain a type of binding and can be knotted with each other in pairs, with a process called as correlation. With the help of entanglement property, the spin of a particle (up or down) can be correlated with the spin of another particle. Because of the phenomenon of superposition, the intended element has no single spin direction before being intended, but is concurrently in both a spin-up and spin-down state. The spin state of the element being intended is determined at the time of measurement, and communicate to the concurrent element, which concurrently assume the opposite spin direction to that of the measured element. This phenomenon is called as "spooky action at a distance", the mechanism, which is not yet be explained by any theory. It purely should be taken as known. The Principle of Quantum Entanglement facilitates qubits that are far by infinite distance, to interact with each other instantly. Whatever may be the distance between the concurrent elements, they will remain in entangled state till they are isolated. The Superposition and Entanglement properties of Quantum computing generate an extremely superior computing supremacy. The 2-bit register of a classical computer stores only one of four the four values (00, 01, 10, or 11) at any given time, while a 2-qubit register in a quantum computer is able to store all four numbers at the same

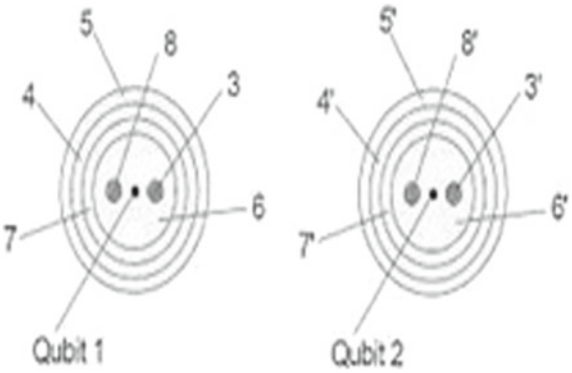


Fig. 3 Entanglement of qubits

time, since each qubit represents two values. In case more qubits are added, then ability of computation will greatly increase, almost exponentially as shown in above Fig. 3.

2.3 Quantum Programming

The quantum computer programming paradigm is somewhat as compared to classical computers. Classical computers basically use instructions to accomplish the task. Based on the number of instructions that executed for some time, the speed and performance of classical computers are decided. On the other hand, the quantum computers follow entirely different logic of programming. Programming environments in quantum computing identifies the number of qubits to be used, and identifying all the relationships, that is possible combinations between them in terms of superposition and entanglement states. By considering all the states which has got from superposition and entanglement the programming will be done for solving problems. It is observed that with a large set of states of qubits, it possible to solve the problem in faster way compared to classical computers.

The Illustrations with quantum programming are as follows.

Illustration 1:

During 1994, Peter Shor, designed a quantum algorithm for the factorization of large numbers. It canters on a system that uses number theory to approximate the consistency of a large number sequence.

Illustration 2:

During 1996, an experiment was conducted, by designing an algorithm that is demonstrated to be the fastest possible algorithm for performing searching operation through unstructured databases. The algorithm was so capable that it was required only, on average, around, M square root, where M is the total number of elements, searches to find the needed output. The same task was performed with classical computer, with an average of $M/2$ searches.

2.4 Quantum Supremacy

Along with the above 2 successful experiments, there is another break through happened in the quantum computing world, i.e. in the area of cryptography. Recently during 2019–20, an experiment was conducted by IBM for testing the performance of classical computer and quantum computer comparatively over the same problem. As a result, it is observed that, classical computer may take approximately 10,000 year or more time to solve the problem, whereas quantum computer solved the problem just in 200 s. This enhances wide opportunities for the researchers in the area of quantum computing. Currently quantum computers are used in very restricted areas, where as the quantum supremacy opened its door wide for developing more quantum algorithms and there by a hope of implementing quantum computers in all areas like business, medical, science etc.

3 Literature Survey

Computers are manufactured for the purpose of computations, in this context the computer era started during nineteenth century. It was the era of vacuum tubes, and other electronic devices which are available in mega sizes. As the development happened in electronic industry, new technologies are introduced in the market, and it produced great impact on the size, performance of classical computers. The evolution of computers can be clearly seen in generation of computers, which depicts the size, capacity, of the classical computers based on technology available in each generation. During 1st Generation the classical computers were developed with the use of vacuum tubes, with few bits of capacity and few hours of processing time, during second generation the classical computers were developed with use of transistors, with few kilo bits of capacity and few minutes of processing speed, the 3rd generation was the era of integrated circuits, where the size and capacity was in few mega/giga bytes with milli/micro second of processing speed, 4th generation was called generation of microprocessors where, capacity of microprocessors were increased immensely, and in 5th generation the scope was on artificial intelligence, where the computers which are used in the field are of size with few tera bytes, with processing speed of nano seconds. Even though classical computers are evolved to a large extent, and are widely used in all most all the fields, still some drawbacks exist while solving some problem with classical logic.

Quantum theories developed during 1900 by Max Planck, in which, the idea was that, the energy exist in individual units called as “quanta”, as the matter does. This theory got developments by a number of scientists over thirty years led to the different way understanding the quantum theory. Based on Quantum Theory, Paul Benioff, proposed a classical computer functioning with some quantum mechanical principles. But it was already accepted that David Deutsch of Oxford University provided the critical impetus for quantum computing research. It was proposed that

using quantum theory a quantum computer can be designed which functions in better way than classical computing.

Bub [1], has demonstrated the Quantum information and computation, where a comparative analysis has been made between classical method and quantum method along with properties. Debnath et al. [2], demonstrated a small programmable quantum computer with atomic qubits. Ashley Montanaro, Quantum algorithms: an overview, which is published as article. The article depicts the simulation environments for quantum computers and how quantum mechanics is used in the construction of quantum computers. El-Din [3], has made a detailed study over Quantum Machine Learning Computation: Algorithms, Challenges, and Opportunities, and as a result the challenges which are there in implementing quantum computing, maintaining the sustainability of the quantum systems under ideal environments are described. About the challenges, the author how demonstrated which are the areas where the quantum technology may be used. Le Gall [4], has demonstrated how Exponential Separation of Quantum and Classical Online Space Complexity can be achieved. Raussendorf and Briegel [5] has demonstrated, A One-Way Quantum Computer, Physical Review Letters. van Dam and Shparlinski [6], designed an approach on Quantum and Classical Algorithms for Exponential Congruences, which states how quantum algorithms are applicable for exponentially varying problems.

In this paper, a comparative study on classical and quantum computers are implemented with illustration on various types of problems. It is observed that some of the problems which are still under discussions are having complexities to solve. To resolve such kind of problems an attempt has made using quantum computing technique. In this research a detailed analysis is made on different applications of quantum computing.

4 Observations

Based on the study made in this paper on classical computer and quantum computer with architectures, methods, and types of problems as examples, the following points are noticed as observations.

It is observed that classical computers are improved tremendously over decades, in terms of computing speeds, and memory sizes. Because of this, classical computers are used in almost all the fields. But still in many fields, classical computers are poor in solving problems in terms of speed, efficiency, accuracy, etc.

Quantum computer is another type of computer considered in this paper and it is observed that quantum computer uses the quantum physics principles for solving the problem. It uses quantum properties such as superposition and entanglement on qubits, thereby creating a different operating principle for solving the problems. It is also noticed by the examples that, quantum computers are superior in solving some problems, where the classical computers are lagging.

5 Future Scope

This paper mainly focuses on a comparative study between classical and quantum computers and it is figured out that quantum computers perform better in many aspects where classical computers are not. By considering the hurdles in development of quantum computers, if quantum computers are developed with sustaining environments, reduced error rates, affordable sizes etc., then quantum computers can be used in all fields such as business, science, etc. In future studies an investigation will be done on the areas where quantum computers are used.

6 Conclusions

In this paper a detailed survey is accomplished on classical and quantum computers. Here working principles of classical and quantum computers are narrated, and a study is made with examples and illustrations, which describes some drawbacks of classical computer and benefits of quantum computer over classical computers. Concerned to the applications of quantum computers, currently it is used in cryptography, teleportation, Non probabilistic problems, high order of growth problems.

References

1. Bub, J. (2005). Quantum information and computation. arXiv preprint quant-ph/0512125.
2. Debnath, S., Linke, N. M., Figgatt, C., Landsman, K. A., Wright, K., & Monroe, C. (2016). Demonstration of a small programmable quantum computer with atomic qubits. arXiv:1603.04512v3 [quant-ph].
3. El-Din, D. M. Quantum machine learning computation: Algorithms, challenges, and opportunities.
4. Le Gall, F. (2008). Exponential separation of quantum and classical online space complexity. arXiv:quant-ph/0606066v3.
5. Raussendorf, R., & Briegel, H. J. (2001). A one-way quantum computer. *Physical Review Letters, 86*(22), 5188–5191.
6. van Dam, W., & Shparlinski, I. E. (2008). An approach on classical and quantum algorithms for exponential congruences. arXiv:0804.1109v1 [quant-ph].

An LSTM-Based Fake News Detection System Using Word Embeddings-Based Feature Extraction

Rishibha Sharma, Vidhi Agarwal, Sushma Sharma, and Meenakshi S. Arya

Abstract Fake news is manipulated news or misinformation that is spread across the Internet with an intention to impose certain ideas and to damage an agency, organization and person often using dishonest, sensationalist and outright fabricated headlines to increase readership. Due to the propagation of fake news, there is a need for computational methods to detect them. Fake news existed for decades, and in the research community, the detection of fake news has been a desired topic. Around 70% of people are concerned about the propagation of fake news. Given the challenges related to the detection of fake news research problems, the researchers globally are trying to figure out the basic attributes of the problem statement. The objective of this paper is to detect whether the online articles are fake or credible, using various machine learning techniques like GloVe word embeddings and long short-term memory (LSTM) as feature extraction and as a classifier technique to find the best fit for the model.

Keywords Fake news detection · NLTK · LSTM · GloVe

1 Introduction

The effect of fake news has become a huge problem as it shapes the views of people. Fake news is spread through various means such as Web sites, social media, articles, etc. that gains credits and helps to manipulate the perception of the readers. In 2017, Tim Berners-Lee, the inventor of the World Wide Web, claimed that fake news was one of the most significant disturbing Internet trends [1]. Through social media platforms like Twitter, Facebook, WhatsApp, it has become easier for wrong content to diffuse quickly. A study conducted by MIT scholars found that fake news stories on Twitter are 70% more likely to be retweeted than true stories. It also takes about six times as long for real news to reach 1500 people as it does for fake reports to reach

R. Sharma · V. Agarwal · S. Sharma · M. S. Arya (✉)
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India
e-mail: raina.arya@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_26

the same number of people [2]. Humans share or retweet false information, which is usually characterized by eye-catching titles or clickbaits. This prevents people from stopping to check and verify the given information. As a result, the news gets circulated without prior checking of the truthfulness of the piece of news.

Content producers and journalists are more motivated toward speed and spectacle over accuracy, and readers often lack the literacy skills required to interpret news critically. Thus, an automated assistive tool that could filter inaccurate information would save a lot of time, effort and help to gain the trust of the readers.

The proposed technique aims to detect whether the online articles are fake or credible, using various machine learning techniques like GloVe word embeddings as feature extraction and long short-term memory (LSTM) as a classifier technique to find the best fit for the model. The organization of this paper is as follows. Related works are reviewed in Sect. 2. Section 3 outlines the proposed theoretical concepts. The proposed models are discussed in Sect. 4. The results of the experiment and discussions are presented in Sect. 5. The paper is summed up and concluded in Sect. 6.

2 Related Works

In [3], a new dataset called "LIAR" has been introduced which is one of the largest datasets in the topic of automatic fake news detection. The datasets of LIAR does not include a full article, and it brings together in total 12,800 manually labeled short bits of statements of the page "polifact.com". They used neural networks to combine meta-data with text, a great improvement to detect fake news.

An elaborated study was examined in [4] using N-gram which was compared with various feature extraction techniques. This paper achieved the highest output while implementing unigram as the extraction method technique and support vector machine (SVM) as the classification model.

In [5], the proposed method investigated fake news in public datasets by using machine learning algorithms like CountVectorizer, TF-IDF Vectorizer, Naive Bayes and NLP. The investigation showed that CountVectorizer was much effective to use as a classifier technique because successfully 89.3% of the news was classified correctly.

A review of identifying false news on social media by focusing on a few attributes that are repeatedly encountered in fake news, along with fake news characterizations based on "social theories and psychology, an existing algorithm from a data mining perspective, evaluation metrics and representative datasets" is presented in [6].

The application of machine learning techniques and natural language processing for detecting fake news accurately is explored in [7]. The paper investigates and compares the accuracy of techniques which are Naive Bayes, support vector machine (SVM), neural network and long short-term memory (LSTM) to find the best fit for the model.

The detection of fake news using unified key sentences can accurately perform sentence matching between article and question by using key sentence retrieval based

on a bilateral multi-perspective matching model presented in [8]. It performed an efficient matching operation for the word vectors obtained using BiLSTM. An accuracy ranging from 64 to 69% has been achieved.

The proposed technique [9] presents a system to detect the stance of headlines about their corresponding article bodies, concentrating largely on the observation of clickbait detection. This method was practiced on publicly known sets of data, attaining a success rate of 89.59%.

3 Theoretical Concepts

3.1 Word Embeddings

Word embedding is a learned depiction of words that have the similar meaning have a very alike representation. It is about improving the ability of networks to learn from text data. By representing that data as lower dimensional vectors. These vectors are called embedding. Independent words are represented as actual esteemed vectors in a predetermined vector space through embedding techniques. Every word is designed to be mapped to one vector, and that vector values are learned to resemble a neural network. Every word is a real-valued vector of 10 or 100's of proportions. The words which are used in the same contexts have similar representations. Word embedding algorithms can figure out tons of relationships from the text data. They use the idea of context and learn by seeing what word occurs near other words (Fig. 1).

GloVe—Global Vectors for Word Representation The proposed technique uses GloVe for word embedding. It is an unsupervised learning technique used for obtaining representations of words in vectors. The approach captures the meaning of one word embedding with the structure of the whole observed data; co-occurrence counts and word frequencies are the main measures. This model trains on the co-occurrence of word counts and makes effective use of statistics by reducing the least-square errors which produce a word vector space with meaningful structure. It produces an outline that preserves the similarities among the words with vector distance.

To store the information, we use co-occurrence matrix X, each entry corresponds to the times the s word exists in the context of the t word.

$$P_{ts} = P(t) = \frac{X_{ts}}{X_t} \tag{1}$$

is the probability the word with s index exists in the context of the t word.

Ratios of probabilities of co-occurrences helps us to begin learning word embeddings. We initially define a function F as follows:

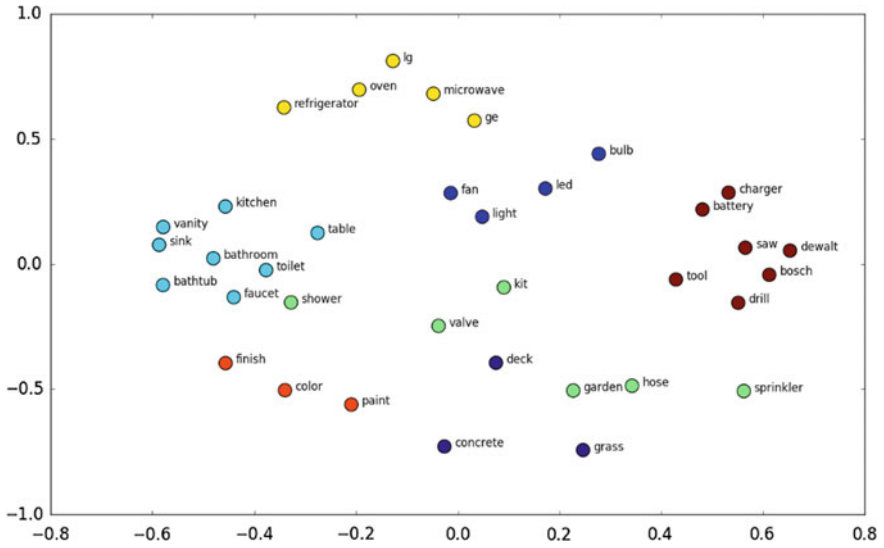


Fig. 1 Vector space of word embeddings. *Source* https://miro.medium.com/max/1050/1*jnrmjaYMF_gq-9zia2HV7g.png

$$F(w_t, w_s, \tilde{w}_k) = \frac{P_{tk}}{P_{sk}} \tag{2}$$

which is dependent on two-word vectors with t and s indexes separating the context vector with k index. F encodes the information which is present in the ratio form; the most inherent way to represent the difference in vector form, i.e., to subtract one vector from the other:

$$F(w_t - w_s, \tilde{w}_k) = \frac{P_{tk}}{P_{sk}} \tag{3}$$

In the equation, the right-hand side is the scalar while the left-hand side is the vector. To avoid this, we can calculate the product of two terms:

$$F\left((w_t - w_s)^T, \tilde{w}_k\right) = \frac{P_{tk}}{P_{sk}} \tag{4}$$

As long as in the co-occurrence of word–word matrix the distinction between standard words and context words is arbitrary, we can replace the probabilities ratio as:

$$F\left((w_t - w_s)^T, \tilde{w}_k\right) = \frac{F\left(w_t^T \tilde{w}_k\right)}{F\left(w_s^T \tilde{w}_k\right)} \tag{5}$$

and solve the equation as follows:

$$F\left(w_t^T \tilde{w}_k\right) = P_{tk} = \frac{X_{tk}}{X_t} \tag{6}$$

Assuming that F function is exp(), the solution becomes:

$$w_t^T \tilde{w}_k = \log X_{tk} - \log X_t \tag{7}$$

The above equation does not preserve the symmetry, so we absorb two of the terms as:

$$w_t^T \tilde{w}_k + b_t + \tilde{b}_k = \log X_{tk} \tag{8}$$

Now the loss function that we are trying to minimize is the linear regression function with certain changes:

$$J = \sum_{ts=1}^{V} f(X_{ts})(w_t^T \tilde{w}_s + b_t + b_s - \log X_{tk})^2 \tag{9}$$

where f is the weighting function.

3.2 Long Short-Term Memory

LSTM, which stands for long short-term memory, is an extension of the previously famous recurrent neural network (RNN). In addition to RNN, LSTM also has memory over the long run. It is inherently nothing but a neural network. To conquer the disadvantages of traditional RNN, three types of gates are attached to the system for an easy notion of memory. At each timestep, an LSTM cell can choose to read, write or reset the cell by using an explicit gating mechanism. Each time the cell reads an input, the memory is kept updated. Figure 2 illustrates LSTMs with four types of gates, namely forget gate (f), input gate (i), memory gate (c) and output gate (o).

Given C_{t-1}(an old memory), C_t (the new cell memory) is computed as:

$$C_t = f_t * C_t - 1 + i_t * C_{et} \tag{10}$$

Forget gate: The forget gate determines the information that is to be removed from the memory. Given an input at x_t and time step t, which is computed as:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{11}$$

C_{t-1} then gets multiplied with this f_t to transform it with the information removed.

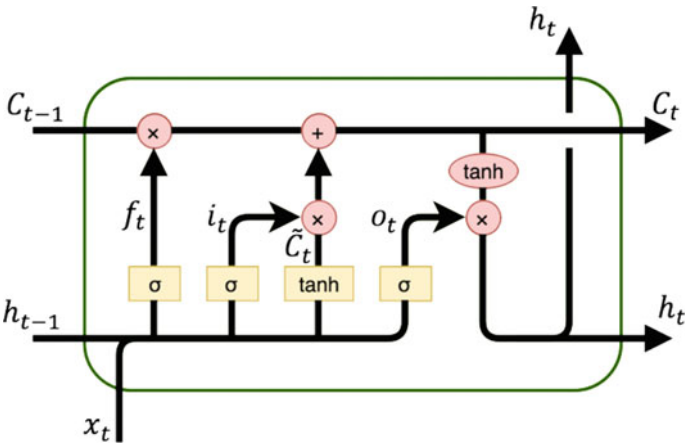


Fig. 2 Internal structure of LSTM networks. *Source* https://apmonitor.com/do/uploads/Main/lstm_node.png

Memory gate: The memory gate is used to generate a new candidate memory. Given an input x_t, which is computed as:

$$C_{et} = \tan h(W_C x_t + U_c h_t - 1 + b_c) \tag{12}$$

Input gate: The input gate determines the amount of information of the candidate memory will be injected into an updated one. Given an input x_t, which is computed as:

$$i_t = \sigma\left(W_f x_t + U_i h_{t-1} + b_i\right) \tag{13}$$

C_{et} is then multiplied by it, and the new memory is added to the cell memory.

Output gate: The output gate is used to determine how much of the memory cell is extracted. It is computed as:

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{14}$$

the new hidden state is then updated as:

$$h_t = o_t * \sigma_c C_t \tag{15}$$

Due to the presence of internal memory and its ability to get updates sequentially, the long dependency problem is resolved.

4 Proposed System

The architecture of the proposed method is described in Fig. 3. The detailed explanation is given below.

Step 1: Preparing the dataset:

- Converting the data to lowercase and removing the punctuations.
- Removing the stop words such as "is", "at", "that", etc. that create noise in the data.
- Tokenizing the data to convert the text into a sequence of integers (each integer is the index of a token in a dictionary).

Step 2: Applying the GloVe Word embedding

- Representing each word in a sentence by a dimensional vector.
- Obtaining the relationship between two words that are not only the same but also syntactic (e.g., "watch" and "see" are syntactically very different in, but their meaning is related and similar).

Step 3: Creating a sequential model

- Creating models using LSTM and embedding layers.
- Feeding the embedding matrix to the embedding layer.
- Defining the LSTM layer which returns the sequence data.
- Training the model
- Detecting whether the news article is fake or not using testing data.

The corpus of data implemented in this project has around 6335 articles of data. The main features included in each row of the data are id, title, text, label, classification of being fake or true [3]. Some examples of the dataset are presented below (Table 1).

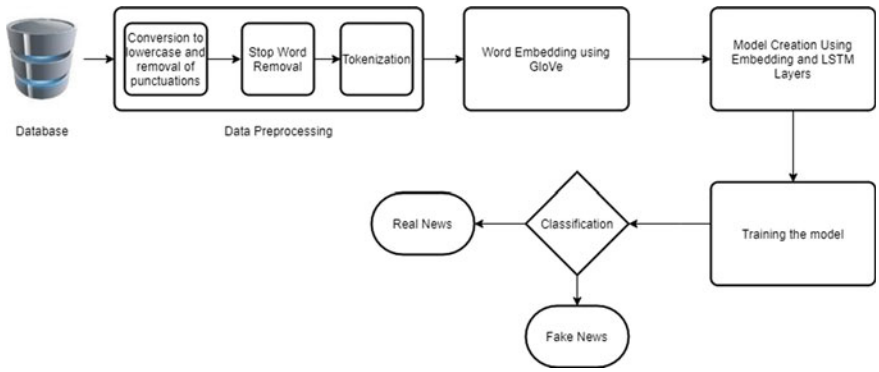


Fig. 3 Block diagram of the proposed system

Table 1 Dataset description

ID	Title	Text	Label
8476	You Can Smell Hillary's Fear	Daniel Greenfield, a Shillman Journalism Fello…	FAKE
10294	Watch The Exact Moment Paul Ryan Committed Pol…	Google Pinterest Digg Linkedin Reddit Stumbleu…	FAKE

5 Results and Discussions

In the proposed technique, GloVe word embedding was applied along with the LSTM classifier to detect fake news. After performing best for our model, our intention was to calculate the closeness of the predicted model to the original one as an evaluation metric. Provided the size of the data, it becomes easy for a recurrent neural network (RNN) model to overfit, i.e., it becomes difficult or unable to predict the accuracy of the unseen data. A dropout layer was added to overcome overfitting and to improve generalization.

The LSTM-based model was trained over 10 epochs, obtaining an accuracy of 0.84 on both the validation and the test set. On increasing the epochs by 10 the training loss decreased and the improvement in accuracy rate was not too significant.

The most accurate results were achieved with a dropout rate of 0.l and 10 epochs. The performance is promising as GloVe word embeddings LSTM achieved an accuracy of 84.1%.

6 Conclusion

Fake news and other types of false information may take on numerous facets. They can also have big impacts, as information forms our view of the world: We make essential knowledge-driven decisions. By using our proposed technique, we would be able to predict if a certain news item is fake or credible, thereby improving the faith of readers and helping them in their decision-making process. Our model utilizes GloVe for implementing word embedding to create an embedding matrix of our dataset which is then passed as an input to our model along with LSTM layers. We then trained our dataset and achieved an accuracy of 84.1%. We had run our model on the "news.csv" dataset. In future, we shall execute this model on additional public datasets like the LIAR dataset. We can further increase model size, tweak the hyperparameters, use bidirectional LSTM and test if the model performs better.

References

1. Pathak, G. Curb Fake News, Not Media. *Telangana Today*, telanganatoday.com/curb-fake-news-not-media.
2. Dizikes, P. (2018) Study: On twitter, false news travels faster than true stories. *MIT News*, 8 Mar. 2018, news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308.
3. Wang, W. Y. (2017). "Liar, Liar Pants on Fire": A new benchmark dataset for fake news detection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers). https://doi.org/10.18653/v1/p17-2067.
4. Ahmed, H., Traore, I., & Saad, S. (2017). Detection of online fake news using N-Gram analysis and machine learning techniques. *Lecture Notes in Computer Science Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, 127–138. 10.1007/978-3-319-69155-8_9.
5. Agudelo, G. E. R., Parra, O. J. S., & Velandia, J. B. (2018). Raising a Model for Fake News Detection Using Machine Learning in Python. *Lecture Notes in Computer Science Challenges and Opportunities in the Digital Era* 596–604. https://doi.org/10.1007/978-3-030-02131-3_52.
6. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media. *ACM SIGKDD Explorations Newsletter, 19*(1), 22–36. https://doi.org/10.1145/3137597.3137600.
7. Prannay, S. R., Diana Elizabeth, R. P., Manoj, M. K., & Poonam V. T. (2019). A study on fake news detection using naïve bayes, SVM, neural networks and LSTM.
8. Kim, N., Seo, D., & Jeong, C.-S. (2018). FAMOUS: Fake news detection model based on unified key sentence information. In *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*. https://doi.org/10.1109/icsess.2018.8663864.
9. Bourgonje, P., Schneider, J. M., & Rehm, G. (2017). From clickbait to fake news detection: An approach based on detecting the stance of headlines to articles. In *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing Meets Journalism*. https://doi.org/10.18653/v1/w17-4215.

Emotion and Depression Detection from Speech

Yash Deshpande, Shreya Patel, Meghan Lendhe, Manpreet Chavan, and Reeta Koshy

Abstract Human speech communication conveys semantic information of the underlying emotions corresponding to the speech of the interlocutor. So, detection of emotions by analysis of speech is important for identifying a subject's emotional state. Numerous features from human speech are used by convolutional neural network (CNN) and support vector machine (SVM) techniques to detect the emotions such as anger, happiness, fear, sadness, surprise and neutral that are associated with the speech. Prolonged sadness is considered the prerequisite for depression. Monitoring subject's speech over a period of time helps in detecting clinical depression. Databases of different accents of English language are taken to make sure the system incorporates multiple accents. Emotion and depression detection have applications in fields like lie detection, military, counseling, database access systems, etc.

Keywords Emotion detection using machine learning · Depression detection · SVM · Supervised neural networks · Mel frequency cepstral coefficients · Health informatics · Spectral feature extraction

Y. Deshpande (✉) · S. Patel · M. Lendhe · M. Chavan · R. Koshy
Sardar Patel Institute of Technology, Mumbai, India
e-mail: yash.deshpande@spit.ac.in

S. Patel
e-mail: shreya.patel@spit.ac.in

M. Lendhe
e-mail: meghan.lendhe@spit.ac.in

M. Chavan
e-mail: manpreet.chavan@spit.ac.in

R. Koshy
e-mail: reeta_koshy@spit.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_27

1 Introduction

Sentiment analysis is one of the most useful and significant applications of machine learning and is used extensively for understanding the sentiments (emotions) from various tweets or comments online. Research shows that a person's voice contains information that reflects the mental state of that person and hence can be used as an indicator for detection of depression and its severity [1]. Also, speech features do not vary broadly, making it easier to record and process as compared to video or text features. Detection of depression (and thereafter seeking therapy) can help improve the subject's mental health and thus become more productive. Emotion detection from speech is a relatively unexplored arena with many interesting possibilities which can have many innovative and a plethora of applications in the near future, especially in the healthcare sector for easy detection of clinical depression. Many new customer services can use speech emotion detection to provide better services. For example, in call centers automated systems can detect specifically angry callers and prioritize their target customers accordingly.

Mental health is a rising matter of concern among the younger generation. Mental well-being is as important as physical well-being. According to a report by the World Health Organization, more than 264 million people suffer from depression worldwide. As depression has become a common cause for increased suicide cases, early detection and its treatment thus become important. But many people suffering from mental disorders often do not open up easily because of the stigma attached to mental health. Currently, most of the depression detection cases are mainly assessed through subjective evaluations of the symptoms by a clinical expert. This method is very labor-intensive and time-consuming. Also, the evaluation is inconsistent subject to the professionals and is expensive [1]. Detection of depression from speech can facilitate the diagnosis and treatment of patients. Machine learning provides an easy and reliable pre-screening tool of detecting severity of depression from speech. Chat bots can be made more humane to evoke appropriate responses considering the mood and emotion of the user. Thus, emotion detection from speech has a lot of potential for use in interactive systems and providing better customer service.

2 Literature Survey

Emotion detection is not restricted to only tweets or online comments, but even speech has enormous scope for its application. Emotions affect human experience, especially leadership skills in a big way. They have a very significant influence on one's life [2]. Identification of emotion from speech signal can aid professionals in counseling students, ensuring proper mental stability of employees [3] and help new businesses grow [4]. Feature extraction is one of the most important steps involved in emotion detection [5].

For detecting emotions, algorithms use prosodic and spectral features. Pitch, inflection, etc., are some types of prosodic features [2]. Spectral features try to decipher frequency segments from speech. Mel-frequency cepstral coefficients (MFCCs) represent audio signals in frequency domain. MFCCs help in studying the manifestation of speech through the vocal track. This helps in converting the continuous speech signal into pseudo-stationary fragments of frequency which can be treated as discrete. MFCCs manipulate bands of frequency in logarithmic fashion which is more accurate for detecting human speech than the linearly spaced frequency bands of DCTs or FFTs [6].

For emotion detection, MFCCs show better performance on well-researched databases like Berlin Database of Emotional Speech (emoDB) than other coefficients [4]. Low-level descriptors can be extracted for working in various domains like space, time, etc. However, global descriptors prove to be more useful of windowed signals. Global descriptors are aggregated to obtain feature vectors.

Classification models can be built using linear discriminant analysis (LDA) and support vector machines. For the emoDB dataset, SVM classifier performed well for one-vs-one classification for most emotions, except for anger and sadness due to similar acoustic signatures. LDA classifier was not significantly accurate due to sparse amount of data. On the other hand, the RED database had 73.3 and 71.8% accuracy for SVM and LDA, respectively. Overall, SVM was better at both multiclass and binary classification. Sonawane et al. [7] used SVM for classification. MFCCs were used for feature extraction. SVM could only classify fixed length data and not variable length data. Both multiple, viz. linear and nonlinear, techniques were used for emotion classification. Better performance was obtained using nonlinear multiple SVM. Multiple SVM as a classifier suffered on a training data set of multiple languages and accents for speech features of MFCC [7].

Harar [8] used deep neural network (DNN)-based approach for emotion detection. EmoDB was used containing 271 labeled recordings but only for three emotions, viz. anger, neutral and sad. The DNN architecture had only fully connected layers—one of size 480 and another of size 240; and a final softmax layer for output. Stochastic gradient descent algorithm was used for training the model. Input data batched in the accuracy obtained using this approach was 96.97%.

Berlin Database of Emotional Speech (BDES) upon choosing different classifiers along with different acoustic features showed variations in the emotions detected. Acoustic features used can be linear prediction cepstral coefficients or Mel-frequency cepstral coefficients. Support vector machine (SVM) classifier showed better accuracy as compared to GMM on BDES considering MFCC as speech features [9]. Parameters of convolutional neural network (CNN) and long short-term memory (LSTM) layers were trained on interactive emotional dyadic motion capture (IEMOCAP). The pace at which the two models trained was significantly different than each other [10]. Gradient boosting on The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) and Surrey Audio-Visual Expressed Emotion (SAVEE) datasets showed that SAVEE has higher accuracy in detecting emotions than RAVDESS. While considering live data, i.e., spontaneous, non-acted, genuine speech samples from subjects, gradient boosting showed better results than

both K-Nearest Neighbors (KNN) and SVM [11]. It was also evident that pitch feature and MFCC feature yielded different accuracy on BDES dataset with HMM classifier [9]. When audio segments were classified independently on BDES using deep neural network's (DNN) stochastic gradient descent, DNN had no knowledge of the actual context of the samples of speech it was supposed to detect emotions from [8].

There exists a strong correlation between the speech features and the presence or absence of depression. This was studied by Sahu et al. [12] by using the Mundt database. While comparing the sustained vowel utterances and free-flowing speech, attributes like jitter, shimmer and degree of breathing were considered to extract MFCC features and train the SVM model for classification. Average magnitude difference function (AMDF) was used to quantify the parameters, and it was observed that shimmer, jitter and breathiness were high for a depressed person. Mari et al. [9] also found that quantifying the jitter and shimmer, degree of breathing computed from an AMDF on Mundt database on Hamilton depression scores showed that dip profiles contain important information about the state of a depressed person.

When a speaker decides to speak something, a series of information is processed in his brain. Thus, her thoughts, feelings and sense of well-being affect her speech. Apart from spectral features, depression affects not only the semantic (content of information) but also syntactic (structure of information) features. Mitra et al. [13] found that spontaneous speech is better than read speech for depression detection and refutes the traditional clinical data collection methods.

Chlasta et al. [14] used the Distress Analysis Interview Corpus (DAIC) dataset with several pre-trained CNN architecture models replacing the last layer of the model as per the dimensions of their dataset. By fine-tuning the parameters, they built a strong classification model with two classes—depressed and non-depressed. Tasnim et al. [15] made an attempt to develop a system that captures the users' voice and analyzes it to detect depression severity. They proposed a comparative study for distinguishing depressed and non-depressed individuals (binary classification problem) and determining severity of depression (a regression problem). They compared four models—random forest model, SVM, gradient boosting tree (GBT) and deep neural network on the AVEC 2013 and AVEC 2017 datasets. It was observed that DNN was the most effective model for classifying the depressed and non-depressed individuals.

The Patient Health Questionnaire depression scale (PHQ-8) scores and Beck scores [16] are some of the popular scales for diagnosing depression. The classification accuracies obtained were similar in both cases. In He et al. [16], AVEC 2013 and AVEC 2014 databases were used for developing DCNN classification model. It asserted that deep-learned features performed better than hand-crafted features (energy descriptors) which are tedious to derive. But the combination of both features boosted the performance.

3 Dataset

For developing an emotion classifier, the datasets which we have used are Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS), Toronto emotional speech set (TESS) and Surrey Audio-Visual Expressed Emotion (SAVEE).

RAVDESS provides audio samples in North American English accent for four emotions that include happy, sad, angry and fearful. RAVDESS consists of 7356 files in three modality formats, i.e., audio-only (16bit, 48 kHz.wav), audio-video (720p H.264, AAC 48 kHz,.mp4) and video-only (no sound) for each of the 24 actors (12 females, 12 males).

The SAVEE database was recorded from four native English male speakers, postgraduate students and researchers at the University of Surrey aged from 27 to 31 years. Emotions for which audio samples are recorded are anger, disgust, fear, happiness, sadness and surprise. A neutral category is also added to provide recordings of seven emotion categories. The text material consisted of three common, two emotion-specific and ten generic sentences that were different for each emotion and phonetically balanced.

In TESS, there are a set of 200 target words spoken by two actresses (aged 26 and 64 years) portraying each of seven emotions (anger, disgust, fear, happiness, pleasant surprise, sadness and neutral). There are 2800 audio files in total. The dataset is organized such that each of the two female actors and their emotions are contained within its own folder. And within that, all 200 target words audio file can be found. The format of the audio file is a WAV format.

For depression detection, the DAIC-WOZ database was used which is part of Distress Analysis Interview Corpus (DAIC) [1]. DAIC-WOZ is a collection of 189 sessions of subjects with extensive questionnaire. A particular session was labeled "depressed" if PHQ8 score for that session was greater than or equal to 10. Since it is a highly imbalanced dataset, around only a quarter of the 189 files are depression-labeled and the rest are non-depression-labeled. Out of 189 files, we used 49 files of depressed subjects and non-depressed subjects each. Furthermore, out of the many types of data available in the corpus for every session, the COVAREP feature files along with the audio files were used (Fig. 1).

4 Methodology

4.1 Feature Extraction

Features for emotion detection were extracted using pyAudioAnalysis [17] python library. All 34 features, viz. zero crossing rate, energy, entropy of energy, spectral centroid, spectral spread, spectral entropy, spectral flux, spectral rolloff, 13 MFCCs, 12 chroma vectors and chroma deviation, were used to train both the models. Window size of 50 ms was used to make frames for extracting features.

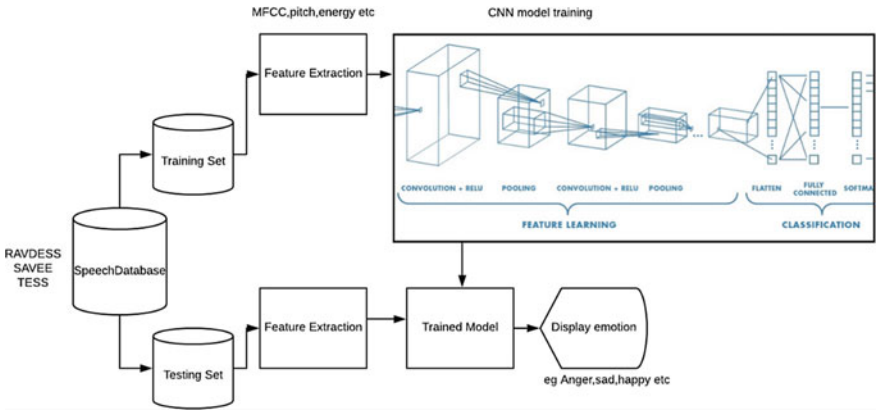


Fig. 1 System architecture

On the other hand, for detecting depression, features extracted using COVAREP [18] were used from DAIC-WOZ dataset. Window size of 10 ms was used to frame the audio signals. A total of 74 features were extracted—the details of which can be found in the documentation.

4.2 Classification Models

In this work, two algorithms—SVM and CNN—are used for emotion classification. The emotion classes used for the classifier were Anger, Disgust, Fear, Happy, Neutral, Sad and Surprise. Four cases were analyzed by training both SVM and CNN on first only on MFCC features and then on other acoustic features as well.

Before training the SVM model, the feature vectors of predictor variables were standardized between 0 and 1 with mean and standard deviation of the data 0 and 1, respectively. The response variable, i.e., emotion classes was label-encoded. Radial basis function (RBF) kernel with gamma kernel coefficient 0.1 was used for SVM classifier for training only MFCC features, while gamma value of 0.01 was used for training all 34 features (Figs. 2 and 3).

The convolutional neural network (CNN), on the other hand, comprised of eight convolutional layers with ReLU activation function, two pooling layers and a dense

$$min_w \lambda \parallel w \parallel^2 + \sum_{i=1}^{n} (1 - y_i \langle x_i, w \rangle)_+$$

Fig. 2 SVM hinge-loss function

$$\frac{\delta}{\delta w_k}\lambda \parallel w \parallel^2 = 2\lambda w_k$$

$$\frac{\delta}{\delta w_k}\left(1 - y_i\langle x_i, w\rangle\right)_+ = \begin{cases} 0, & \text{if } y_i\langle x_i, w\rangle \geq 1 \\ -y_i x_{ik}, & \text{else} \end{cases}$$

Fig. 3 SVM gradients

layer with softmax activation function. This configuration was specifically used as it gave the most stable model. The training and testing losses for the said model neatly converged. Adam optimizer was used for training the network with 0.01 as learning rate. Both only MFCC features and the other 34 features were trained on the same network.

For detecting depression, a 2-D CNN was used with four convolutional layers each with a pooling layer and a global average pooling layer with batch size of 32 for training.

5 Experimentation and Results

Figure 4 shows the accuracies for the two models—SVM and CNN when trained on only MFCCs and other acoustic features along with MFCCs. In both the cases, SVM gave better accuracy compared to CNN as SVM works better in classification problems especially in emotion recognition.

While extra features made the accuracy of SVM increase by almost 4%, CNN accuracy dipped by around 14%. While SVM found a better hyperplane to classify

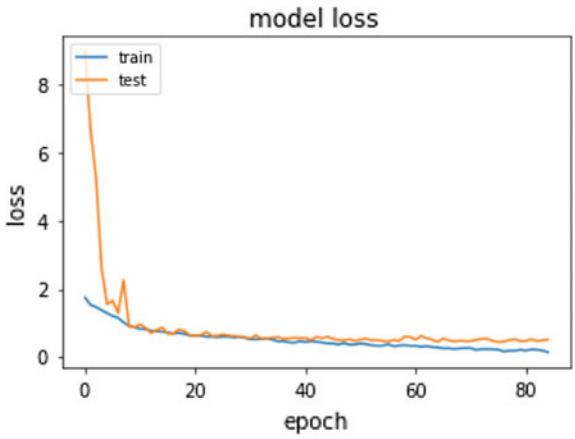


Fig. 4 Loss with MFCC only

Table 1 Accuracies of emotion detection algorithms

Algorithm	MFCCs	MFCCs + Others
SVM	84.53	88.26
CNN	84.21	70.31

data points with extra features, CNN model could not make use of the increased dimensions.

In light of an imbalanced and scarce dataset and computational limitations of our machines, depression accuracy with 2-D CNN remained at 60% (Table 1).

6 Conclusion

The model detects different emotions from speech, and its applications range from counseling, health care to music recommendation. But most of the previous work only incorporates one kind of accent. Hence, pan accent work has not been established yet in emotion detection. This thesis attempts to overcome this issue by incorporating speech data from various accents to provide a more comprehensive emotion detection system. We have also found that neural network worked as good as SVM classifier on only one type of features but as the dimension of data increased, CNN model already struggling with small size of data could not find relevant patterns and its accuracy fell behind SVM's.

In future, the accuracy can further be increased with intelligent feature selection. This work can be coupled with a real-time application for both emotion and depression detection that can process the spontaneous speech instead of the acted emotions.

References

1. Gratch, J., Artstein, R., Lucas, G., Stratou, G., Scherer, S., Nazarian, A., Wood, R., Boberg, J., DeVault, D., Marsella, S., Traum, D., Rizzo, S., & Morency, L. -P. (2014) The distress analysis interview corpus of human and computer interviews. In *Proceedings of Language Resources and Evaluation Conference (LREC)*.
2. Rajvanshi, K. (2018). An Efficient approach for emotion detection from speech using neural networks. *International Journal for Research in Applied Science and Engineering Technology*, *6*, 1062–1065. https://doi.org/10.22214/ijraset.2018.5170.
3. Sharma, U. (2016). Identification of emotion from speech signal. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi* (pp. 2805–2807).
4. Kamaruddin, N., Abdul Rahman, A. W., & Abdullah, N. S. (2014). Speech emotion identification analysis based on different spectral feature extraction methods. In *The 5th International Conference on Information and Communication Technology for The Muslim World (ICT4M), Kuching* (pp. 1–5). https://doi.org/10.1109/ict4m.2014.7020588.

5. Semwal, N., Kumar, A., & Narayanan, S. (2017). Automatic speech emotion detection system using multi-domain acoustic feature selection and classification models. In *2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), New Delhi* (pp. 1–6). https://doi.org/10.1109/isba.2017.7947681.
6. Chamoli, A., Semwal, A., & Saikia, N. (2017). Detection of emotion in analysis of speech using linear predictive coding techniques (L.P.C). In *2017 International Conference on Inventive Systems and Control (ICISC), Coimbatore* (pp. 1–4). https://doi.org/10.1109/icisc.2017.8068642.
7. Sonawane, A., Inamdar, M. U., & Bhangale, K. B. (2017) Sound based human emotion recognition using MFCC & multiple SVM. In *2017 International Conference on Information, Communication, Instrumentation and Control (ICICIC), Indore* (pp. 1–4). https://doi.org/10.1109/icomicon.2017.8279046.
8. Harár, P., Burget, R., & Dutta, M. K. (2017). Speech emotion recognition with deep learning. In *2017 4th International Conference on Signal Processing and Integrated Networks (SPIN), Noida* (pp. 137–140). https://doi.org/10.1109/spin.2017.8049931.
9. Lukose, S., & Upadhya, S. S. (2017). Music player based on emotion recognition of voice signals. In *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, (pp. 1751–1754). https://doi.org/10.1109/icicict1.2017.8342835.
10. Sudhakar, R. S., & Anil, M. C. (2015). Analysis of speech features for emotion detection: A review. In *2015 International Conference on Computing Communication Control and Automation, Pune* (pp. 661–664). https://doi.org/10.1109/iccubea.2015.135.
11. Iqbal, A., & Barua, K. (2019) A real-time emotion recognition from speech using gradient boosting. In *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 1–5) https://doi.org/10.1109/ecace.2019.8679271.
12. Saurabh, S., & Carol, E. -W. (2016). Speech features for depression detection. (pp. 1928–1932). https://doi.org/10.21437/interspeech.2016-1566.
13. Mitra, V., & Shriberg, E. (2015). Effects of feature type, learning algorithm and speaking style for depression detection from speech. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* pp. 4774–4778. https://doi.org/10.1109/icassp.2015.7178877.
14. Chlasta, K., Wołk, K., & Krejtz, I. (2019). Automated speech-based screening of depression using deep convolutional neural networks. In *8th International Conference on Health and Social Care Information Systems and Technologies*, Oct 2019. https://doi.org/10.1016/j.procs.2019.12.228.
15. Tasnim, M., & Stroulia, E. (2019). *Detecting Depression from Voice.* Cham: Springer Nature Switzerland AG 2019, April 2019, (pp. 472–478). https://doi.org/10.1007/978-3-030-18305-9_47.
16. He, L., & Cao, C. (2018). Automated depression analysis using convolutional neural network from speech. *Journal of Biomedical Informatics, 83,* 103–111.
17. Giannakopoulos, T. (2015). pyAudioAnalysis: An open-source python library for audio signal analysis. *PLoS ONE, 10*(12), e0144610. https://doi.org/10.1371/journal.pone.0144610.
18. Degottex, G., Kane, J., Drugman, T., Raitio, T., & Scherer, S. (2014). COVAREP—A collaborative voice analysis repository for speech technologies. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy 2014.

Comparative Study on Parkinson Disease Dignosis Treatment Classification Using Machine Learning Classifier (PDMLC)

Hiral R. Patel, Ajay M. Patel, and Satyen M. Parikh

Abstract Parkinson's infection is a cerebrum issue that prompts shaking, solidness, and trouble with strolling, parity, and coordination. Parkinson's side effects for the most part start step by step and deteriorate after some time. As the infection advances, individuals may experience issues strolling and talking. Parkinson disease is a neurological issue, which is one of the most excruciating, risky, and nondurable infections, which happens at more established ages. The Static Spiral Test (SST), Dynamic Spiral Test (DST) and Circular Motion Test(CMT) on certain point records were utilized in the analytical modelling application which was created for the conclusion of this malady. The dataset is collected from the online authorized source to classify the test basis on the symptoms. These datasets were separated into 70–30% of splitting into the training and testing information inside the original and control structures of ten-fold cross-validation approval method. Training set is utilized for train the ML-based analytical classification models such as logistic regression, neighbors classifier, decision tree, and support vector machine. Model is also trained by performing hyper-parameter tuning. After training phase, the model is assessed with test and the comparative results are discussed. Additionally, new information investigation was completed. As indicated by the outcomes acquired, SVM with RBF kernel is more effective than other classifier; furthermore, logistic regression calculations are in investigation of new information. This study contributes in the same direction by analyzing the behavior according treatment classification.

Keywords Circular motion test · Decision tree · Dynamic spiral test · KNN · Logistic regression · Parkinson diseases · Static spiral test · Support vector machine

H. R. Patel (✉) · A. M. Patel · S. M. Parikh
Faculty of Computer Applications, Ganpat University, Gujarat, India
e-mail: Hrp02@ganpatuniversity.ac.in

A. M. Patel
e-mail: Ajaykumar.patel@ganpatuniversity.ac.in

S. M. Parikh
e-mail: Satyen.parikh@ganpatuniversity.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_28

1 Introduction (Heading 1)

Parkinson's sickness is one of the most well-known development issue on the planet. Parkinson's influences roughly 1% of all grown-ups beyond 60 years old. This issue is the consequence of the degeneration of dopamineproducing cells in the substantia nigra in the midbrain. Dopaminergic meds are presently being utilized as a treatment to improve a significant number of the manifestations that portray Parkinson's sickness. The reason for the ailment cannot known. Be that as it may, there has been a bounty of research attempting to distinguish potential inceptions. This writing survey will feature a portion of the principle places that analysts have on the reasons for Parkinson's illness. Among these causes are natural poisons, hereditary variables, and oxidative pressure. Parkinson's malady (PD) is the second most regular development ailment after basic tremor with a rate pace of around 1% among grown-ups beyond 60 years old. In spite of the fact that the specific reason for Parkinson's ailment cannot known, it is considered to be potentiated by the cooperation of natural, furthermore, hereditary components [1].

At the point when a patient presents with tremor, it very well may be helpful to play out a couple of basic pen and paper tests. Right now, disclose how to expand the benefit of penmanship and of drawing Archimedes spirals and straight lines as clinical evaluations. These errands take merely seconds to finish however give an abundance of data that supplements the standard physical assessment. They help the determination of a tremor issue and can add to its longitudinal observing. Viewing the patient's upper appendage while they compose and draw may uncover variations from the norm, for example, bradykinesia, dystonic acting, and distractibility. The completed content and drawings would then be able to be assessed for recurrence, plentifulness, bearing and balance of oscillatory pen developments and for in general size of handwriting. Fundamental, dystonic, practical, and parkinsonian tremor every ha a trademark example of variation from the norm on these pen and paper tests [2].

Spiral drawing test identifies indications of Parkinson's in patient. A test that includes drawing a winding on a piece of paper could be utilized to analyze early Parkinson's ailment. Australian specialists have trialed programming that estimates composing rate and pen pressure on the page. Both are valuable for identifying the malady, which causes shaking and muscle inflexibility. The Melbourne group said the test could be utilized by GPs to screen their patients after middle age and to screen the impact of medicines. The investigation, distributed in Frontiers of Neurology, included 55 individuals—27 had Parkinson's and 28 did not. Speed of composing and pen pressure while portraying are lower among Parkinson's patients, especially those with a serious type of the ailment. [BBC Health News Report for RMIT on 6 September 2017]. The subsequent test is the Dynamic Spiral Test (DST). Not at all like SST, Archimedean spiral just shows up and vanishes in certain time interims, at the end of the day the Archimedean winding flickers. This powers the patient to remember the example and keep on drawing. Another test Circular Motion test is

Only participants with PD completed the repetitive finger movement task such as circle drawing, in persons with Parkinson's disease (PD) which impact fine motor tasks may improve the assessment of persons with PD [3].

2 Literature Survey

The earlier research work states that, huge possibilities for studies are available in PD recognition concentrated on discourse preparing [4–6] where the conclusion is finished utilizing continued vowels and characteristic discourse. Moreover, engine side effects can additionally be distinguished and managed, displaying persistent developments and walk [7, 8]. Change in the kinematics of penmanship is one of the underlying signs saw in PD. McLennan et al. [9] detailed that roughly 5% of patients with PD indicated micrographia (anomalous little letter size) and 30% of patients announced declining of penmanship before the beginning of engine manifestations. Engine manifestations related with Parkinson's infection (stiness, bradykinesia, and tremor) cause three fundamental changes recorded as a hard copy [10]: the size of composing [11] (micrographia [12]), pen pressure [13], also, kinematics. A few devices have been created to dissect PD quiet penmanship [14]. Not just the static angles, yet in addition, the dynamic ones are intriguing, for example, speed, and pen pressure decrease during composing [13, 15]. A few survey papers have been distributed as of late [16, 17]. The penmanship of an individual relies upon the visual ability [12], composing style, or language aptitudes of the individual [18], indicating an enormous between subject inconstancy. An option in contrast to penmanship is the utilization of drawings.

The basic strategy for defining the diseases stages is as effectively expressed in the basic part, and there are numerous prospects to group or subtype PD: period of beginning, clinical phenotypes (engine and non-engine), infection seriousness, or neuropathological changes. Author presents the current logical information on as often as possible utilized orders and along these lines need to give the premise to resulting stratification of patients, that is expected to distinguish the ideal treatment for the individual patient. Authors recognize that a characterization or arranging of PD as a heterogeneous neurodegenerative ailment is somewhat fake, yet at the same time, think of it as significant, particularly considering the developing exceptionally explicit causative treatment ideas [19].

An assorted scope of AI classifiers was analyzed to locate the most noteworthy all out exactness for PD analysis. The decision tree and bolster vector machine classifiers are developed with using Scikit-Learn an AI library [20]. Models are upgraded using stratified cross approval with exactness, F-1, review, and accuracy as measurements. The dataset can be ordered using a decision tree classifiers' progression. It includes standard decision trees, arbitrary woods, slope supported decision trees, and additional tree classifiers also.

A decision tree works classifying the information homogeneously between the two classes [23, 24]. In total, these detachments make an arrangement precision over

the preparing set that is the applied to the testing set to evaluate speculation. Unique subsets of information model are prepared and used for decision tree model. Irregular timberlands can be utilized to develop unique subsets of information [25]. Additionally, these models are also applied for precision for tests not utilized in the sub-trees and parameters are tuned to boost the normal exactness of the model over the preparation set. Tree classifiers [26] are also another variety of decision tree classifiers. It relays on stochastic strategies that create comparatively shallow however more extensive decision trees. Slope-supported decision trees work by making straightforward poor classifiers that isolate the example space. The poor classifiers are consolidated to limit a differentiable misfortune work [27]. The calculation iteratively changes the past arrangement state by making another classifier for the preparation set. This procedure is rehashed to create an outfit of classifiers that can arrange the preparation set precisely.

3 Methodology

In this paper, the comparative study is carried out for comparing the results of different classifiers with parameter tuning and discuss the results. The classifier models are implemented in Python using sklearn library.

There are different classifier is applied on the parkinson dataset.

a. **Logistic Regression**: Logistic regression(LR) is a machine learning calculation which is to calculate a probabilty of a certain event or class, i.e., Yes/No, Positive/Negative. It uses the likelihood concept and is a prescient investigation calculation. Consider a logistic regression, a linear regression model yet the logistic regression utilizes an increasingly mind boggling cost work, this cost capacity can be characterized as the 'Sigmoid capacity' or otherwise called the 'strategic capacity' rather than a direct capacity tends it to constrain. The cost work lies in the range of 0 and 1 and this boundary restrictions are defined by the theory of strategic relapse. In this way, direct capacities neglect to speak to it as it can have a worth more prominent than 1 or under 0 which is beyond the realm of imagination according to the speculation of calculated relapse.

$$\mathrm{O} \leq h_\theta(x) \leq 1$$

Logistic regression hypothesis expectation

b. **K-Nearest Neighbor (K-NN)**: k-nearest neighbors (k-NN) is a regulated AI model. Administered learning is the point at which a model gains from information that is as of now named. An administered learning model takes in a lot of info articles and yield esteems. The model at that point prepares on that information to figure out how to outline contributions to the ideal yield so it can figure out how to make expectations on concealed information. k-NN models work by taking an information point and taking a gander at the 'k' nearest marked information focuses. The information point is then alloted the name of most of the 'k' nearest

focuses. The kNN undertaking can be separated into composing three essential capacities:

- Figure the separation between any two focuses
- Discover the closest neighbors dependent on these pairwise separations
- Larger part vote on a class marks dependent on the closest neighbor list

c. **Decision Tree**: DS is one of tree-based classifier which generates the rule-based tree by using major two criterian such as gini index or entropy. In this paper, both cases for DS is implemented.

$$\text{GiniIndex} = 1 - \sum_{J} P_j^2$$

if a random variable x can take N different value, the ith value x_i with probability $p(x_i)$, we can associate the following entropy with x:

$$H(x) = -\sum_{i=1}^{N} p(x_i) \log_2 p(x_i)$$

d. **Support Vector Machine (SVM)**: Support vector machine (SVM) is famous and amazing classifier technique. SVMs, much like strategic relapse, it builds an ideal isolating hyperplane in the component space between the two classes. SVMs are capable to precisely perform non-direct grouping by means of the bit stunt and this is SVMs main advantage. The part stunt ventures the information into a higher measurement, where it turns out to be directly detachable. Fitting a SVM includes hypertuning parameters X and γ. γ decides the impact of information focuses, higher qualities make the model increasingly worldwide, and low qualities mean information focuses influence a littler neighborhood gathering. X is the regularization parameter, which directs the smoothness of the model. Low estimations of X relate to a smoother and less complex model, but the probability of misclassified information focuses is higher. Higher estimations of X will precisely group more information focuses by expanding the multifaceted nature of the model. So, in this work, the SVM is utilized by implementing the model using different types of kernel.

The basic model representation of this work is represented in Fig. 1.
The following are the stages of the model implementation:

- Data gathering [1 lakh plus instances are collected through the online-authenticated data source.
- Data pre-processing [Missing Value handling, feature extruction, data scaling, and encoding]
- Data spliting [Using ten-fold cross-validation, the 70–30% data split among train-test dataset]

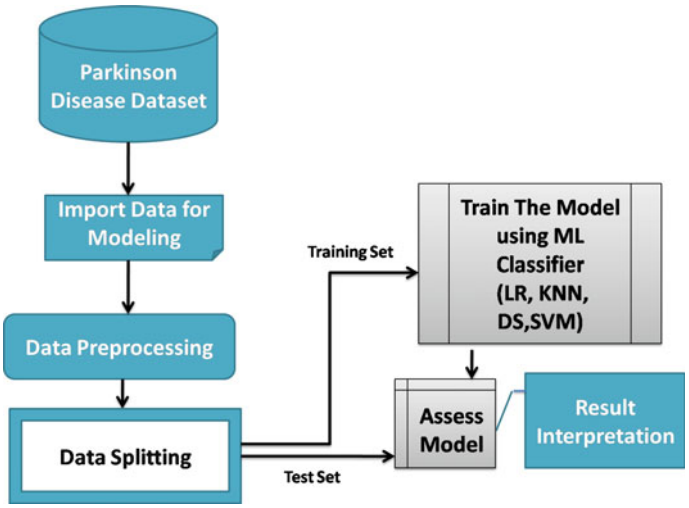


Fig. 1 Proposed model

- Model implementation using different classifiers with parameter tuning using train set
- Assess models using test dataset
- Discuss the results by considering the classifier model performance measures.

4 Implementation and Result Discussion

The model is implemented with four different classifiers. The confution matrix, accuracy score, and classification report(F-Measures, Recall, and Precision) are main performance measurement for the classification type of analytical problem in machine learning. There is a lot of disarray in applied ML about what an approval dataset is actually and how it varies from a test dataset. Accuracy is the thing that we generally mean, when we utilize the term exactness. It is the proportion of number of right expectations to the all out number of information tests.

$$\text{Accuracy} = \frac{\text{Number of Correct predictions}}{\text{Total number of predictions made}}$$

A confusion matrix (CM) is a synopsis of expectation results on an order issue. The quantity of right and inaccurate expectations are outlined with tally esteems and segregate in different class. The disarray lattice shows the manners by which your order model is befuddled when it makes expectations. It gives us knowledge not just into the mistakes being made by a classifier however more significantly the sorts of blunders that are being made (Table 1).

Table 1 Confusion matrix [4–6]

Confusion Matrix	Class Type 1 Predicted	Class Type 2 Predicted
Class Type 1 Actual	True Positive (TP)	False Negative (FN)
Class Type 2 Actual	False Positive (FP)	True Negative (TN)

Here, as per this work, the Class Types are 3 as per below [4].

- 0: Static spiral test (Draw on the given spiral pattern)
- 1: Dynamic spiral test (Spiral pattern will blink in a certain time, so subjects need to continue on their draw)
- 2: Circular motion test (Subjectd draw circles around the red point)

So, originally, accuracy is nothing but the

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Recall(R): Review can be characterized as the proportion of the complete number of effectively ordered positive models gap to the absolute number of positive models. High Recall demonstrates the class is effectively perceived (few FN).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Precision(P): To get the estimation of accuracy we separate the complete number of accurately grouped positive models by the all out number of anticipated positive models. High Precision demonstrates a model marked as positive is without a doubt positive (few FP).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

High review, low precision: This implies that the vast majority of the positive models are effectively perceived (low FN) yet there are a ton of bogus positives.

Low R, High P: This shows that we miss a ton of positive models (high FN); however, those we foresee as positive are in reality positive (low FP).

F-measure: Since we have two measures (Precision and Recall), it assists with having an estimation that speaks to them two. We compute a F-measure which utilizes harmonic mean instead of arithmetic mean as it rebuffs the extraordinary qualities more. The F-Measure will consistently be closer to the littler estimation of Precision or Recall.

The following image shows the characteristics of the predictor (Fig. 2).

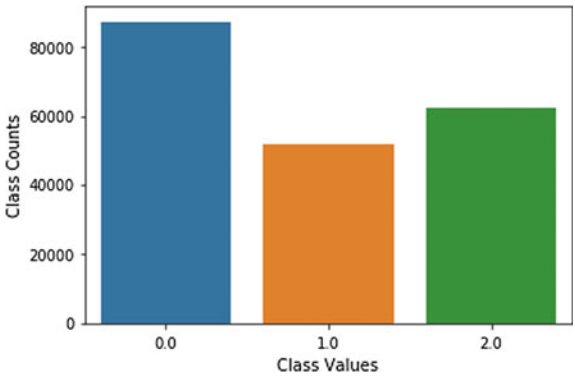


Fig. 2 Predictor characteristics [4–6]

$$\text{F - measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

The proposed work modelling implementation is carried out using Python Sci-Kit Learn library. The below table shows the implementation results by performing parameter tuning (Table 2).

As per the experimental study, by performing different classifiers, the decision tree gives highest accuracy rate with stable modelling score. So the following tree is generated as result using DST with gini criterion applied (Fig. 3).

5 Conclusion

The Parkinson disease classification is carried out by applying automated architect of machine learning classifier. These results are promising because without manipulating the data and techniques, the results are incorporated. Now, in the view of point classifier, the decision tree gives best fitness model when we apply without altering the depth of tree and selecting gini as splitting criterion. The SVM model is compromised with accuracy when sigmoid as kernel is applied due to the non-binary classification problem. As per the comparison study, it may be possible to get more optimized result by applying the machine learning's best classifier neural network. So for, the future study of the case study will discuss the results by applying NN.

Table 2 Result comparision using classification performance measures [4–6]

Classifier name	Parameter tunned	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
Logistic regression (LR)	C = 1.0, fit_intercept = True, random_state = None, solver = 'liblinear',	62	62	62	61
K-nearest neighbor (KNN)	algorithm{'auto', 'ball_tree', 'kd_tree', 'brute'}(Model tune four way but the results are same. weights = 'uniform'	72	73	73	73
Decision tree (DT)	Criterion='gini'	94	96	94	93
	Criterion='entropy'	93	95	93	92
	Criterion='gini', max_depth = 4	74	76	74	73
	Criterion='entropy' max_depth = 4	73	75	73	72
Support vector machine (SVM) İn Python (SVC)	kernel = 'linear'	61	61	61	61
	kernel = 'poly'	73	73	73	73
	kernel = 'rbf'	75	75	75	75
	kernel = 'sigmoid' [Due to non-binary class accuracy suffered]	53	53	53	53

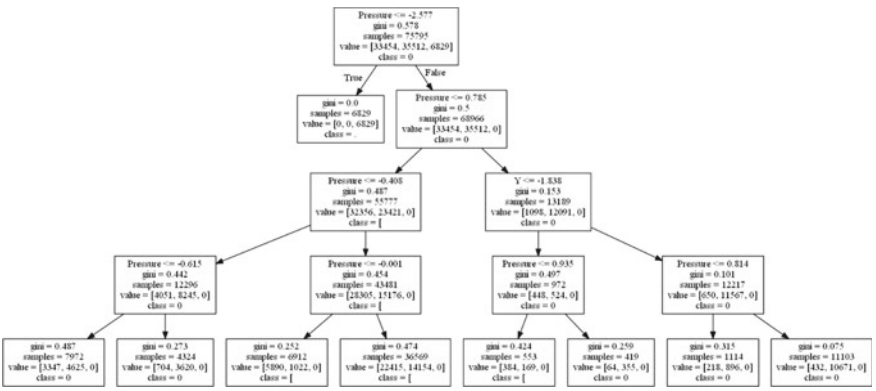


Fig. 3 DST with criterion gini for splitting [4–6]

References

1. Shafique, H., Blagrove, A., Chung, A., & Logendrarajah, R. (2011). Causes of parkinson's disease: Literature review. *Research and Reviews in Parkinsonism., 1,* 5–7. https://doi.org/10.2147/JPRLS.S37041.
2. Zham, P., Kumar, D. K., Dabnichki, P., Arjunan, S. P., Raghav, S. (2017). Distinguishing different stages of parkinson's disease using composite index of speed and pen-pressure of sketching a spiral. Front Neurol. 2017; 8: 435. Published online 2017 Sep 6. https://doi.org/10.3389/fneur.2017.00435 PMCID: PMC5592741.
3. Stegemöller, E. L., Zaman, A., & Uzochukwu, J. (2019). Repetitive finger movement and circle drawing in persons with Parkinson's disease. *PLoS ONE, 14*(9), e0222862. https://doi.org/10.1371/journal.pone.0222862.
4. Lahmiri, S.; Shmuel, A. Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. Biomed. Signal Process. Control 2018, 49, 427–433. [CrossRef].
5. Gómez-García, J.A.; Moro-Velázquez, L.; Godino-Llorente, J.I. On the design of automatic voice condition analysis systems. Part I: Review of concepts and an insight to the state of the art. Biomed. Signal Process. Control 2019, 51, 181–199. [CrossRef].
6. Gómez-García, J.A.; Moro-Velázquez, L.; Godino-Llorente, J.I. On the design of automatic voice condition analysis systems. Part II: Review of speaker recognition techniques and study on the e_ects of di_erent variability factors. Biomed. Signal Process. Control 2019, 48, 128–143. [CrossRef].
7. Viteckova, S.; Kutilek, P.; Svoboda, Z.; Krupicka, R.; Kauler, J.; Szabo, Z. Gait symmetry measures: A review of current and prospective methods. Biomed. Signal Process. Control 2018, 42, 89–100. [CrossRef].
8. San-Segundo, R.; Navarro-Hellín, H.; Torres-Sánchez, R.; Hodgins, J.; De la Torre, F. Increasing Robustness in the Detection of Freezing of Gait in Parkinson's Disease. Electronics 2019, 8, 119. [CrossRef].
9. McLennan, J.E.; Nakano, K.; Tyler, H.R.; Schwab, R.S. Micrographia in Parkinson's disease. J Neurol. Sci. 1972, 15, 141–152. [CrossRef].
10. Zham, P.; Arjunan, S.P.; Raghav, S.; Kumar, D.K. E_cacy of Guided Spiral Drawing in the Classification of Parkinson's Disease. IEEE J. Biomed. Health Inform. 2018, 22, 1648–1652. [CrossRef].
11. Potgieser, A.R.; Roosma, E.; Beudel, M.; de Jong, B.M. The e_ect of visual feedback on writing size in parkinson's disease. Parkinsons Dis. 2015, 2015, 857041. [CrossRef].
12. Drotár, P.; Mekyska, J.; Rektorová, I.; Masarová, L.; Smékal, Z.; Faundez-Zanuy, M. Evaluation of handwriting kinematics and pressure for di_erential diagnosis of Parkinson's disease. Artif. Intell. Med. 2016, 67, 39–46. [CrossRef].
13. Letanneux, A.; Danna, J.; Velay, J.L.; Viallet, F.; Pinto, S. From micrographia to Parkinson's disease dysgraphia. Mov. Disord. 2014, 29, 1467–1475. [CrossRef].
14. Thomas, M.; Lenka, A.; Kumar Pal, P. Handwriting Analysis in Parkinson's Disease: Current Status and Future Directions. Mov. Disord. Clin. Pract. 2017, 4, 806–818. [CrossRef].
15. Rosenblum, S.; Samuel, M.; Zlotnik, S.; Erikh, I.; Schlesinger, I. Handwriting as an objective tool for Parkinson's disease diagnosis. J. Neurol. 2013, 260, 2357–2361. [CrossRef].
16. Impedovo, D.; Pirlo, G. Dynamic Handwriting Analysis for the Assessment of Neurodegenerative Diseases: A Pattern Recognition Perspective. IEEE Rev. Biomed. 2019, 12, 209–220. [CrossRef].
17. Impedovo, D.; Pirlo, G. Chapter 7: Online Handwriting Analysis for the Assessment of Alzheimer's Disease and Parkinson's Disease: Overview and Experimental Investigation. In Series on Language Processing, Pattern.
18. Recognition, and Intelligent Systems. Pattern Recognition and Artificial Intelligence; World Scientific Publishing: Singapore, 2019; pp. 113–128.

19. Classification of advanced stages of Parkinson's disease: translation into stratified treatments Rejko Krüger, Jochen Klucken, Daniel Weiss, Lars Tönges, Pierre Kolber, Stefan Unterecker, Michael Lorrain, Horst Baas, Thomas Müller, Peter Riederer J Neural Transm (Vienna) 2017; 124(8): 1015–1027. Published online 2017 Mar 24. doi:https://doi.org/10.1007/s00702-017-1707-x.
20. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg et al.,"Scikit-learn: Machine learning in python," Journal of machine learning research, vol. 12, no. Oct, pp. 2825–2830, 2011.
21. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin,S. Ghemawat, G. Irving, M. Isard et al., "Tensorflow: A system for large-scale machine learning." in OSDI, vol. 16, 2016, pp. 265–283.
22. F. Chollet et al., "Keras," 2015.
23. Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics, 21*(3), 660–674.
24. L. Breiman, Classification and regression trees. Routledge, 2017.
25. Liaw, A., Wiener, M., et al. (2002). Classification and regression by randomforest. *R news, 2*(3), 18–22.
26. Simm, J., de Abril, I. M., & Sugiyama, M. (2014). Tree-based ensemble multi-task learning method for classification and regression. *IEICE TRANSACTIONS on Information and Systems, 97*(6), 677–1681.
27. J. H. Friedman, "Greedy function approximation: a gradient boosting machine," Annals of statistics, pp. 1189–1232, 2001.

Development of Security Clustering Process for Big Data in Cloud

M. R. Shrihari, T. N. Manjunath, R. A. Archana, and Ravindra S. Hegadi

Abstract Big data is a large volume of data that demands effective cost and novel forms of data to process for enhanced security imminent innovation, assessment creation, and optimization process. Cloud computing provides a consistent, fault lenient, and extensible background to the big data disseminated managing systems. Cloud computing is emerged that has a popular computing model of system of servers hosted over the Internet for the store, manage, and process the data. It includes several advantageous characteristics of cloud storage are easy to access, scalability, resilience, cost efficiency, and reliability of the data. Since every organization is moving its data to the cloud and uses the storage services provided by the cloud. However, threats to cloud security such as data loss, data breaches, insecure application programming interface, account or service traffic hijacking, denial of service, malicious insiders, shared technologies and dangers, insufficient diligence are moving the security of the data. So, it is essential to protect the information against denial of services, unauthorized access or modification of data. The proposed method provides the vital security for different environment to store and access the information file in big data cloud. A numerous levels of security services such as authorization, authentication, data integrity, and confidentiality are ensure by applying the proposed method; therefore, only the genuine users are allowable to store and access the data through this method. The proposed method has multiple levels of security such as authorization by the Cloud Service Provider, encryption of the data using

M. R. Shrihari (✉) · T. N. Manjunath
Department of ISE, BMS Institute of Technology and Management, VTU, Karnataka Bengaluru, India
e-mail: shrihari.mr@gmail.com

T. N. Manjunath
e-mail: manju.tn@gmail.com

R. A. Archana
Department of CSE SVIT, VTU, Karnataka Bengaluru, India
e-mail: archana.tnm@gmail.com

R. S. Hegadi
Department of Computer Science, Central University of Karnataka, Kalaburagi, India
e-mail: rshehadi@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_29

different cryptographic algorithms though uploading the data to the cloud, and usage of one time password for decrypting the data while downloading it from the big data cloud.

Keywords Big data · Security cloud service provider · Secure dynamic bit standard

1 Introduction

Cloud computing has many applications such as enabling access to expensive applications at no cost, reducing both establishment and running expenses of computers and software as there is no need for any foundation. Users can place the data at any place. All users are required to associate with a system, say the Internet. Cloud computing in progress as a tool intended for interpersonal computing but currently it is extensively employ for admission software online, online storage space exclusive of tormenting concerning communications cost and processing power. Institute can delegate their Information Technology (IT) infrastructure in the cloud and gain access. Not only private organizations are moving to cloud computing, but the government is also moving some parts of its IT infrastructure to cloud. Big data comprises the digital data from several digital sources which contain sensors, scanners, statistical modeling, videos, cellular phones, digitizers, Internet, e-mails, and public networks which are increasing the data rate. Recent advances in big data cloud computing have accommodated many opportunities for researchers on cryptography to improve the security of authentication and authorization for efficiently accessing the stored data [1].

Big data cloud computing turns into a useful and mainstream business model because of its appealing components. In accumulation to the reimbursement at dispense, the previous components additionally result in genuine cloud-particular security issues. The general populations concern is the security in the cloud, and the users are delaying to exchange their business to cloud. Security issues have been the hindrance to the improvement and broad utilization of cloud computing. Understanding the security and protection chances in cloud computing, creating rich and robust solutions are necessary for its prosperity. In spite of the fact that clouds enable customers to stay away from start-up costs, diminish operating costs, and increment their speed by quickly getting services and infrastructural resources when required. In advertising and business, most of the industries use big data, but the fundamental properties of security may not be implemented [2]. If a safety contravene happen to big data, it would consequence in still more severe authorized consequence and reputational damage than at present. In this novel era, a lot of companionship is with the expertise to accumulate and evaluate petabytes of data concerning their companionship, industry, and their clients. For creation big data protected, method such as encryption, classification, and honey pot discovery should be essential. In various institute, the employment of big data for fraud recognition is extremely smart and

constructive. The argument of safety in cloud computing environments can be classified into network intensity, user authentication intensity, data intensity, and generic concern [3].

2 Multilevel Security System Using SDBS Algorithm

Big data cloud computing properties are distributed through on demand services for consumers over the Internet. Whenever information is shared using cloud by any business or government agency or an individual, there exists a doubt on its security and confidentiality. Data security in a cloud is interrelated to data user authentication. The organization will have its own security policies such as some of the employees are not permitted to access a certain amount of data. Data integrity is also a major part of information security system. This service protects data from unauthorized modification or deletion. The objective of work is to store the data securely and have a more safe and secure access to the data. An efficient security system provides four-level securities which are used for multiple users who can access data in the cloud [4]. Using Secure Dynamic Bit Standard (SDBS) algorithms, the data provider encrypts the data, and it is uploaded to the cloud. The data provider uploads the secured data, and Cloud Service Provider (CSP) permits the data user to access the data. Then, this permitted person is known as the authorized person who can access the data. While downloading the data, data user enters the One Time Password (OTP) to get the data downloaded and decrypted using SDBS algorithm [5].

3 Secure Dynamic Bit Standards (SDBS)

Secure Dynamic Bit Standard (SDBS) have 3-bit levels such as 128 bits, 256 bits, and 512 bits. Whenever the data provider wants to upload a data to the cloud any one of the bit levels is selected randomly, and it will get converted into bytes. Based on byte value, Cloud Service Provider (CSP) will generate the master key and a session key using a random generator [6]. The master key is encrypted by the session key, and both the encrypted master key and the session key will be sent to the data provider. The session key will be used to decrypt the master key, and the master key is used to encrypt the data. The encrypted data will be uploaded to the big data cloud server along with the Proof of Ownership (PoW) which is generated by CSP [7].

If a data user wants to download a data from the big data cloud server, by request, then the CSP will send the encrypted master key and the session key along with the encrypted data to the data user after One Time Password (OTP) verification. Then, using this session key, the master key will be decrypted, and using the master key, the data will be decrypted and stored in the system of the data user. SDBS algorithm is a novel algorithm which has three various standards with eight rounds for 128-bit keys, 10 rounds for 256-bit keys, and 12 rounds for 512-bit keys [3].

4 SDBS Encryption Process

The substitution of bytes is a nonlinear exchange of bytes that functions self-sufficiently on every byte of the condition using a exchange bench called S-box.

4.1 ShiftRows

The ShiftRows is a mechanism of operating on the rows of the state and present it as a cyclic shift that the bytes are cyclically shifted in each row by an individual offset. In this SDBS, for 128 bit standard 4×4 bytes of 8 bits are represented and shifted. The row of first each and every byte of the row will be getting shifted one by one to the left. Likewise, the second and third rows will get shifted by two bytes and three bytes of rows, respectively. The block of sizes in the SDBS has three levels of bit standards, they are 128 bits, 256 bits, and 512 bits of the shifting pattern will be the same [8].

4.2 Mix Columns

Mix column transformation is one of the complex strategies used. Here, a replacement has been done in SDBS by Dynamic Mix Column Transformation (DMCT). A DMCT comprises the multiplication of dynamic matrices which is based on the default matrix known as a matrix of SDBS [9].

4.3 AddRound Key

The addition of round key is the state of combining the SubKeys with the output of Mix column state. In each round, each SubKeys will be derived from the main key using the key schedule algorithm. All the Subkeys in this state will be of the same size. These Subkeys were added by the combination of each byte in the state according to the byte of the SubKeys using bitwise Exclusive OR (XOR) [8].

4.4 Key Schedule Algorithm

Key schedule algorithm is used to produce SubKeys for the various rounds. The round stable is a word in which the three rightmost bytes are constantly 0. Consequently, the effect of Exclusive OR (XOR) of a word is to solitary achieve an XOR on the

leftmost byte of the word. Automatically, random bit standards (128, 256 and 512) will be allocated for each file. Cloud service provider will generate a random master key and session key which is based on random bit level selected [10].

5 SDBS Decryption Process

In SDBS, decryption process the data provider must login and then select a data, and the data user need to download. Before downloading the data, the data user sends the request to the Cloud Service Provider to get master key and session key. Cloud Service Provider encrypts the master key with the session key and sends to the data user. One Time Password is also sent by the Cloud Service Provider to the data users email id or mobile number. If the entered one time password is valid; then, the encrypted data will be downloaded from the cloud, and the decrypted master key with the session key is used to start the decryption process. Finally, the decrypted file is stored into the system [11].

5.1 AddRound Key Transformations

In the backward add round key transformation, called inverse AddRoundKey, the 128 bits of state are bitwise XORed with the 128 bits of the round key. The operation is done reversely as a column-wise operation between the 4 bytes of a state column and one word of the round key [9].

5.2 Inverse ShiftRows Transformation

In the bytes, in the first three rows will get moved circularly right by the 4 bytes are shifted depending on the row number. This reverse function of shift rows will be accessed by using InvShiftRows() function. The row of first every byte of the row will be getting shifted one by one to the right [8].

5.3 Inverse SubstitutionBytes Transformation

InvSubBytes() is the inverse of the byte exchange conversion; the inverse substitution (S)-Box is used for each byte. Every row and column of each byte substituted a state of S-Box reversely [8].

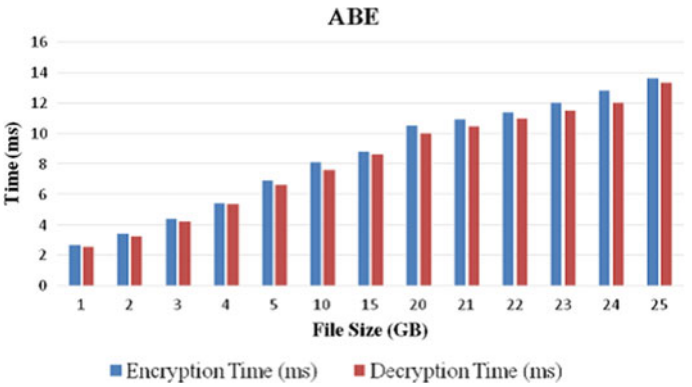


Fig. 1 Performance assessment of system 1—ABE

6 Results and Discussion

The complete investigation is based on four methods, primarily the Attribute-Based Encryption (ABE) was executed; then, the Advanced Encryption Standard (AES) was executed; the third method hybridizes the accomplishment of both ABE and AES; lastly, the Secure Dynamic Bit Standard (SDBS) was executed during the identical situation.

6.1 Performance Assessment of System 1—ABE

The organization systems enclose public key for encryption process and private key for decryption process, so it is called asymmetric algorithm. An attribute-based encryption method typically consists of a key authority, sender, and recipient, generate public and private keys and subject the keys to sender and recipient.

Then (see Fig. 1) represented the performance assessment of system 1—Attribute-Based Encryption (ABE). This routine is conceded out between encryption time and decryption time vs. file size. The 1 GB data took 2.69 ms for encryption, and the 24 GB data took 12.8 ms for encryption. In the similar way, the decryption time is compared with the file size. From the Fig. 1, 1 GB data file took 2.56 ms for decryption, and the 24 GB data took 11.98 ms for decryption [10].

6.2 Performance Analysis of System 2— AES

Advanced encryption standard is a block cipher having default block size of 128 bits as well as allows three a variety of key lengths such as 128, 192, or 256 bits. The

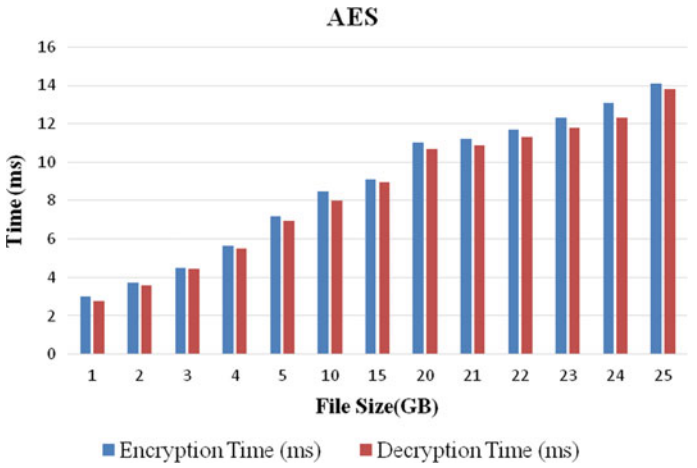


Fig. 2 Performance assessment of system 2—AES

encryption technique encompass of 10 rounds of dispensation 128-bit keys. The 4 × 4 matrix of bytes is used in 128-bit block.

Then (see Fig. 2) depicts the performance of method 2. The assessment is conceded out between encryption time and decryption vs. file size. Since the (see Fig. 2), it is obvious that when the file sizes enhance, and its corresponding encryption time are also increased. So the encryption and decryption instance of method 2 is high when compare with the technique 1. The 1 GB data took 2.99 ms for encryption; also the 24 GB data took 13.1 ms for encryption. In the similar method, the decryption time is evaluated with the file size. From the Fig. 2, 1 GB data file took 2.76 ms for decryption; also the 24 GB data took 12.3 ms for decryption [12].

6.3 Performance Assessment of System 4—Secure Dynamic Bit Standard

The performance of Secure Dynamic Bit Standard (SDBS) algorithm might be study by two kinds of constraint which are encryption time and decryption time. Encryption time is the quantity of time essential to absolute the encryption process equivalent to the file size, and the decryption time is the amount of time required to complete the decryption process consequent to the file size [13].

On top of (see Fig. 3) describes the performance of encryption time and decryption time vs. file size in system 4. The graph carried out between encryption time and file size. The encryption or decryption time of the proposed system 4 is less time when compared with the other techniques system 1, system 2, and system 3. In this technique took 2.56 ms for encrypting 1 GB data; also the 24 GB data file took

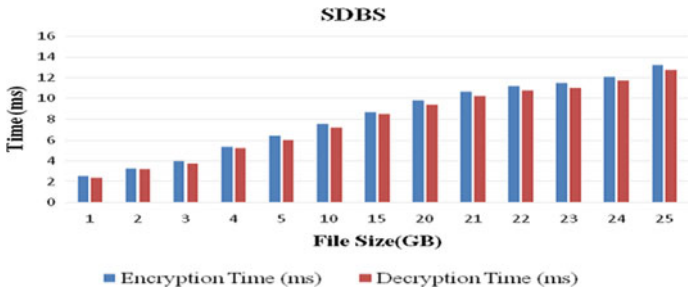


Fig. 3 Performances of system 4—SDBS 128-Bits standard

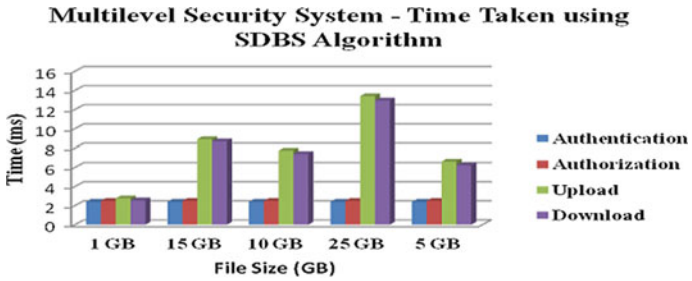


Fig. 4 Multilevel security system—time taken using SDBS 128-bits standard algorithm

12.1 ms for encryption. 1 GB data file took 2.35 ms for decryption; also the 24 GB data file took 11.72 ms for decryption [12].

6.4 Proposed Security Levels

The planned work with system 4—SDBS algorithm has security when compared with the other system like system 1, system 2, and system 3. Multilevel security of authentication, authorization, upload, and download time are considered in milliseconds is shown below (see Fig. 4). The overall scheme run by the scheme is intended with dissimilar organizer sizes vs. time calculated for every step is shown the (see Fig. 4).

7 Conclusion

This investigate has planned a protected method for big data cloud user. The data will be uploaded to the cloud only after authenticated and authorized by the Cloud Service Provider. In a big data cloud situation, the data is stored in an encrypted

method. In the preliminary stage, data supplier desires to inventory for upload the data into the big data cloud. Subsequent to implementation the registration procedure, the data user needs to login to upload the data along with proof of ownership. Based on the credentials, the Cloud Service Provider provides the upload privileges to the data user. The planned scheme has data confidentiality and high data integrity. Through upload or download process, the data are encrypted or decrypted using the proposed SDBS algorithm. This narrative SDBS algorithm proposed in this proposal supply high-level security of the accumulate data. For the moment, it enhance the security against the cloud assault such us involvement Attack, Brute Force Attack, and Structured Query Language Injection Attack. It takes less time to encrypt and decrypt the data. So, once compare with the dissimilar algorithms the planned SDBS technique supply high security for the data. Big data cloud storage aptitude and relevance are explicated. The potential of dissimilar attacks are illustrated in detail. In this proposal, data transmit to the big data cloud has been done by the data supplier. The planned method has high data reliability and data storage exclusive of data loss. Earlier to the data has been uploaded into the storage area; a high-secure algorithm called SDBS is been used. Big data examination information it assists to recognize every data supplier and data user usage of file size, encryption time or decryption time, and upload time or download time.

References

1. Soumyashree, & Nithyashree, K. (2007). A survey on security and privacy methods of bigdata in cloud computing environment. *International Journal of Advance Research, Ideas and Innovations in Technology, 3*(6), 1499–1504.
2. Yu, Y., Au, M. H., Ateniese, G., Huang, X., Susilo, W., Dai, Y., et al. (2017). Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage. *IEEE Transactions on IF&S, 12*(4), 767–778.
3. Zhang, X., Tang, Y., Wang, H., et al. (2019). Lattice-based proxy-oriented identity-based encryption with keyword search for cloud storage. *Information Sciences, 494,* 193–207.
4. Chase, J., Niyato, D., Wang, P., Chaisiri, S. & Ko, R. (2017). A scalable approach to joint cyber insurance and security-as-a-service provisioning in cloud computing. *IEEE Transactions on Dependable and Secure Computing.*
5. Nahar, A. K., Mongia, K., & Kumari, K. (2018). Cloud Computing and Cloud Security. *IJRAE&T, 4*(1), 1–8.
6. Dai, X., Dai, H., Yang, G., et al. (2019). An efficient and dynamic semantic-aware multi-keyword ranked searchscheme over encrypted cloud data. *IEEE Access, 7,* 142855–142865.
7. Shrihari, M. R., Archana, R. A., Manjunath, T. N., & Hegadi, R. S. (2018). A Review on Different Methods to Protect Big Data Sets", 2018, issue-12 & page-4.
8. Fu, Z., Xia, L., Sun, X., Liu, A. X., & Xie, G. (2018). Semanticaware searching over encrypted data for cloud computing. *IEEE Transactions on Information Forensics and Security, 99,* 1.
9. Yao, J., Zheng, Y., Wang, C., & Gui, X. (2018). Enabling search over encrypted cloud data with concealed searchpattern. IEEE Access.
10. Uma, G., & Jayasimman, L. (2018). Enhanced convergent encryption key generation for secured data deduplication in cloud storage. *Journal of Physics: Conference Series Conference Series, 1142,* 012006.

11. Raviteja Kanakala, V., Chanda Shekhar, K., & Vijay Kumar, G. (2017). Security issues in cloud computing. *International Journal of Advanced Research in Computer Engineering & Technology*, *6*(13, 4), 504–509.
12. Zhang, Y., Chen, X., Li, J., Wong, D. S., Li, H., & You, I. (2017). Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing. *Information Sciences, 379,* 42–61.
13. Ghanbarpour, A., & Naderi, H. (2020). An attribute-specific ranking method based on language models for keywordsearch over graphs. *IEEE Transactions on Knowledge and Data Engineering, 32*(1), 12–25.

Agent Score-Based Intelligent Incident Allocation Engine

Mythili Krishnan, Madhan Kumar Srinivasan, and Chriss Acca Mathews

Abstract This paper demonstrates the ticket allocation problem in an IT operations incident management environment, by infusing a data-driven approach using data analytics and rule-based machine learning techniques. While conventional methods such as cherry-picking of tickets, round-robin methods exhibit bias in terms of the perceived ability of human agents, an analytically suave data-driven approach eliminates such bias, while delivering substantial performance benefits. The ticket allocation problem deals not only with how many tickets to allocate to each resource but also what job to allocate to which resource. Thus, solving two potential riddles baffling the Operations Managers, i.e., allocation as well as assignment. In turn, our solution optimizes the queue waiting time which can reduce the cost and increase customer satisfaction. The solution to the ticket-queuing problem can particularly help in ticket prioritization. Thus, we have dissected the ticket allocation problem in three parts and have attempted to provide solutions to all the three aspects. We assimilated and fused the three-pronged solution into a single solution that can be readily used not only in IT operations but in any operations environment where ticket allocation needs to be optimized. In this paper, we are solving the real-time ticket allocation problem, through agent-scoring approach. This depends on agent success rate and agent effort hours. This is an instantaneous solution that will work in real time, as soon as a ticket lands on the floor.

Keywords Ticket allocation · Analytics · Data analysis · Incident management · Machine learning · Empirical analysis · Agent scoring · Real-time scoring

M. Krishnan (✉) · M. K. Srinivasan · C. A. Mathews
Accenture, Bangalore, India
e-mail: mythili.krishnan@accenture.com

M. K. Srinivasan
e-mail: madhan.srinivasan@accenture.com

C. A. Mathews
e-mail: chriss.acca.mathews@accenture.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_30

1 Introduction

"Incident management" is a process in IT service management that helps in restoring the normal operations, when disrupted, due to repeated issues or major incidents. Incident management is an important part of every business organization because it helps to ensure smooth functioning of day-to-day operations, high level of service quality, availability of network, servers, and applications. In day-to-day operations of IT service desk, there is a substantial dependency on queue managers and agents to seamlessly run the IT operations processes. While some of these operations or tasks are trivial and some are not easily automated, the introduction of analytics and robotic process automation (RPA) solutions can help in garnering substantial time and cost savings. At ITSM or IT service management, the aim is to return to normal conditions caused by disruptions due to incidents with minimal impact on business and operations. Incident is defined as an unexpected event that can temporarily disrupt or cripple the smooth running of operations by causing failures in applications, servers, or databases. Simple login issues across multiple users to an application or server can turn into a major incident thereby making a large-scale impact. There can also be complex incidents that might require considerable amount of research, domain, and technical knowledge to resolve, and there can be repetitive issues/incidents that can be caused by faulty configurations, system upgrades, patching changes, etc. This leads us to the point that that there can be different skill-set requirement for resolution of different kinds of incidents and factors such as SME knowledge, domain expertise, years of experience, familiarity with certain tools, and technologies. There can be a myriad categories of incidents that need to be resolved on time and real-time ticket allocation to the right agents is therefore of utmost importance. The usual process that is followed in an IT operation for incident management is as follows:

1. Incident identification
2. Incident logging
3. Incident categorization
4. Incident prioritization
5. Incident investigation and diagnosis
6. Incident resolution and closure.

Usually, after an incident has been logged, there are ITSM tools that will automatically generate the incidents and send the tickets to the queues of the incident management team. When the incident arrives in the queue, for further investigation and resolution, the incidents are assigned randomly to the available agents on the floor. Herein, comes the need for an efficient and data-driven approach for ticket assignment and allocation that would eliminate human judgement, reduce resolution time, and further reduce the effort of queue managers, so that they can focus on other tasks. Our paper focuses on this methodology with emphasis on ticket prioritization, allocation, and assignment to the most efficient agent with the help of agent scores in real time.

2 Business Problem

A glimpse of different categories of incidents and the support provided would point to the fact that the journey of an incident ticket can traverse several cycles. Overall in terms of priority, the incidents are categorized into high, medium, and low incident tickets. In general, high priority incidents are those that have a large-scale impact and include breakdown of an important application/server that can have major business impact in terms of financial, legal, or regulatory frameworks. Medium priority incidents can affect an individual user like standard IT issues and has minimal financial impact while low priority incidents can involve minor IT fixes. In addition, we can also look at the following impact v/s urgency matrix in Table 1.

The priority response code according to the matrix in Table 1 can be illustrated as in Table 2.

As an example, if the impact of an incident is high and the urgency is also defined as high, then the priority of the incident will be Priority 1, whereas if the impact of an incident is low and the urgency is defined as high, then the priority of the incident will be Priority 3. In this way, the impact/urgency matrix can be used to define the priority of tickets for actioning them based on different business objectives. We will be using this concept in Sect. 4. Now, to discuss the broader business problem, we will be bifurcating the problem into three parts:

- Real-time incident allocation
- Incident prioritization
- Incident assignment.

The business problem here is not only to determine the right agent to whom the incident ticket needs to be assigned to but also to determine how to prioritize the incidents and assign them to the right queue. In a service desk environment, any

Table 1 Incident impact versus urgency matrix

Impact			
Urgency	High	Medium	Low
High	1	2	3
Medium	2	3	4
Low	3	4	4

Table 2 Incident priority response code

Priority of ticket	Target response time	Description
1	Immediate	Critical
2	0–10 min	High
3	0–1 h	Medium
4	0–4 h	Low

disruption to the quality of IT service is termed as an incident. Broadly, once an incident is raised, it is categorized, then prioritized based on the SLA requirements, and finally assigned to an agent. Based on the priority assigned to the ticket and ticket type, the agent then identifies the steps for resolution, any need for it to be escalated or transfers it for updates by external vendors, etc. There are numerous commercial offerings that are used for monitoring IT systems, such as IBM Tivoli, and HP OpenView, to name a few. Such monitoring tools help monitor different aspects of service management and raise alerts, if necessary. Such system generated alerts are often the bulk contributors to the total incident ticket volume. Common system generated alerts pertain to memory issues, CPU utilization, and/or server down among others.

In the recent past, there has been multiple attempts at creating automatic ticket resolution solutions. However, this is an extremely complex task that requires very extensive domain knowledge in various IT applications and technologies. Text mining algorithms employed to categorize tickets, based on the keywords are majorly successful at categorizing system generated tickets but is not particularly useful on user-generated tickets. Post-categorization, the tickets may be resolved by either using the corpus of prior resolutions for similar tickets or by an robotic process automation (RPA) process that has been built.

The level of automation is dependent on the category or type of tickets. The work of Zhou, Tang, Zeng, Shwartz Grabarnik [1] details how incident ticket resolution can be aided by recommendations in an automated services environment. By utilizing historical data, their K-nearest neighborhood (KNN)-based recommender logic uses topic-level feature extraction to provide recommendations. However, they only focus on resolution and the types of resolution irrespective of whoever is involved in the resolution, and they do not focus on the allocation problem. There can be different resolution methods based on an agent's efficiency, tenure, experience, skill and agent-scoring methods can help in addressing these aspects of the allocation problem. On the other hand, there has been relatively lesser work on the actual ticket allocation mechanism. The often-seen service desk approach of allocation of tickets by a queue is often referred to as the push model, where the ticket is pushed to an agent. The disadvantage of the push model is that the decision of the queue manager has a bias arising from his perception of the agent's ability. Further, this model also assumes that agents are identical and that processing times are known; this is far from the truth. Deshpande, Garg, and Suri [2] have approached allocation as a variation of the online scheduling problem. They solve the allocation problem with the help of an auction-based model, where an agent shows interest to work on a ticket and bids for it. While this approach eliminates bias of queue managers in allocation and empowers agents to select tickets based on their competence, it creates heavy dependence on the agent's motivation to bid. In addition, such an approach will inevitably increase the total ticket handling time by the increase in actual allocation time.

3 Related Work

Though the number of papers that discuss about the allocation for human agents is few, there has been quite some work done on online scheduling on identical parallel machines. Processing time for jobs in this classical case is known, and the objective is to minimize completion time or make span. Numerous algorithms have been suggested notably, such as list-processing algorithm [3] and competitive deterministic algorithm [4]. These were generalized by Karger, Philips, and Torng [5], and later improved by Albers [6]. Further work includes, the introduction of eligibility constraints—such as each job is to be processed on only one machine [7]. While all these research papers assume that the processing time is known, Li and Huang [8] looked at online scheduling for jobs with arbitrary release times, i.e., the scheduler does not have all the information about the jobs at the onset, but rather receives it piece by piece.

The challenge with workforce management is to balance SLAs in the presence of agents with different skill levels, tenure, and most often random incoming volumes of tickets/cases for smaller time periods. The inherent complexity of managing support is why often businesses wish to simulate support organization before actual implementation. SYMIAN [9] is one of the many tools that were built for this purpose. However, enumeration is often resorted to, for solving the issue of how many temporary and permanent operators must be present, subject to service-level agreements. Bhandari, Harchol-Balter, and Scheller-Wolf [10] introduced modified-balance-equations-disjunctive-constraints (MBEDC) algorithm for the problem; in the process, they factor in additional variables such as time-variant ticket/customer arrival rate and SLAs and determine when the temporary workers have to be called in.

4 Our Approach

As illustrated earlier, operations teams on the floor work in groups, with specific groups handling incident tickets of a type or category. While there might be overlaps of agents in the purview of cross-training [training to equip the agents to handle incident tickets from more than one category/type], in most cases, agents tend to be part of a single group. In IT operations language, these groups are called assignment groups. When a user logs in an incident ticket, they are required to fill in several mandatory fields in the form to log the incident, such as priority, short description of the issue faced, and broad category the issue falls into. When the user chooses the type/category of the issue faced, often available as a drop-down list, the ticket gets routed to the concerned assignment group. For example, the team that handles server issues may be different from the one that handles issues pertaining to an application. Tagging the appropriate category for the ticket enables the right assignment group to swiftly get into the task of resolving it. Resolution time for the tickets vary,

based on the ticket issue. While simple transactional issues may be completed in 5–10 min, more complex troubleshooting issues might take upwards of 4 hours for complete resolution. While it is possible that a less informed user might choose a wrong type/category, for enumeration purposes, we will keep this as beyond the ambit of our current discussion. In addition to the type/category of a ticket, agent tenure/experience is another feature that can significantly influence the resolution time. Often, less-tenured agents will be given simple tickets/tickets of lower priority for resolution; harder/complex tickets are routed to them only after they have gained substantial exposure to the process.

Often, tenured agents, by their in-depth exposure to the process, have an absolute advantage in the resolution of all types of tickets because of their experience and knowledge of solving different types of incidents over time. Hence, rather than assigning simple tickets to them to capitalize on their shorter resolution time, they are often given queue manager roles or other responsibilities in addition to resolving complex tickets. Once the incident ticket falls into the common/specific queue, the queue manager goes through the incident ticket description and makes an estimate about the time required for completion. The queue manager allocates the incident ticket based on his perceived capability of the agent to handle the issue and the agent's workload. In this way, a queue manager spends about 1–4 min per incident ticket for allocation, on an average. The caveat of allotting incident tickets to agents based on perceived capability is that it might result in bias, based on pre-conceived notions of agent performance.

In this paper, we are proposing to do away with such judgement-based allocation and to use data to identify the best agent to resolve specific types of tickets. We propose data-driven allocation to be based on an "agent score," which in turn is a function of the 'Effort Hours' and the "success rate of an agent." To discuss the larger concept of ticket allocation, we would like to introduce the term "Effort Hours" and "agent success rate." Effort Hour is an estimate of the time that an agent will take to resolve a ticket, and agent success rate is a measure of the effectiveness of the resolution, i.e., a proxy for the ability of an agent to efficiently solve a ticket. Agent score is used to measure the efficacy of agents in resolving an incident ticket type which will be calculated as a function of Effort Hours and agent success rate. Using the data around resolution time, agent success rates and agent workload can help identify the most optimal ticket allocation strategy, thereby removing the need for SME judgement and dependency.

Our methodology can help IT operations managers and incident management team by:

- Improving the resolution time
- Timely escalation of tickets that are high priority/close to missing SLA/pending in queue for a long duration
- Improving agent productivity.

The scores can be built on rolling incident data of at least three months. We will be demonstrating the methodology with mathematical formulas that we have devised with empirical examples.

List of assumptions and notations:

1. G_i: refers to assigned group i, we can have 'i' groups, G_1, G_2,….,G_i where $0 < i < N$ (N: set of natural numbers).
2. X_i: incident ticket count for group G_i
3. T_i: total ticket count group of G_i for, i.e., sum (incident tickets, change requests, and service requests)
4. 'p': working days in a month
5. 'q': working hours a day
6. $U\%$: that effective utilization of an agent
7. V_{ij}: is the volume handled by an agent j of tickets that belong to group G_i
8. X_i: is the total volume that belongs to group G_i
9. (EH_{ij}): Effort Hours, the time taken by an agent j to resolve one ticket belonging to a specific group i
10. ASR_{ij}: agent success rate [ASR] of an agent, and j is a measure of how effective he/she has been in meeting the overall SLA targets
11. Z_i: is the number of tickets that have met SLA
12. S_j: agent score, which is a measure of the efficiency of an agent in resolving an incident ticket type, can be a function of the Effort Hours (EH_{ij}) and agent success rate (ASR_{ij}).

There can be more than one assignment group, i.e., groups where particular type of incidents get assigned to. For e.g., incidents related to Citrix will be assigned to "Citrix assignment group" while login issues related to Windows will get assigned to "Windows group." Hence, we can have 'i' groups, G_1, G_2, …., G_i where $0 < i < N$ (N: set of natural numbers). X_i is the incident ticket count for group G_i. T_i is the total ticket count group of G_i for i.e., sum (incident tickets, change requests, and service requests).

An agent resolves tickets related to incidents, service request, as well as change requests. Since, >90% of the total work is related to incident ticket resolution, we will focus on the methodology for incident allocation, and the same can be extended for service requests and change requests.

$$\text{Percentage of only incident tickets in } G_i[\%Y_i] = \frac{X_i}{T_i}, i \in N, i > 0 \qquad (1)$$

We now must estimate the time spent by an agent working on the tickets. An agent works 'p' days in a month and 'q' hours a day and that effective utilization of an agent is $U\%$.

Effective utilization measure is taken, so that we can account for time spent on training, breaks, holidays taken, i.e., internal and external shrinkages. For e.g., if an agent works for 10 hours and spends 2 hours in lunch break, internal team meeting, etc., then his effective/productive hours is 8 hours. Hence, his effective utilization will be 80%. Utilization targets are normally in the range of 70–80% as industry benchmarks. Hence, effective work hours in a day for an agent can be defined as follows:

$$\text{Effective Work Hours in a Day(EWH)} = q * U\% \tag{2}$$

Total effective work hours spent (per day) working on incident tickets from group "*i*" can be calculated as

$$\text{TSD}_i = \%\, Y_i * \text{EWH} = \%\, Y_i * q * U\% * \tag{3}$$

Total effective work hours spent (per month) working on incident tickets from group "*i*" can be calculated as

$$\text{TSM}_i = \text{TSD}_i * p \tag{4}$$

where p is the total workdays per month

V_{ij} is the volume handled by an agent j of tickets that belong to group G_i, i.e.,

$$\sum_{i>0}^{i} \sum_{j>0}^{j} \left(V_{ij}\right) = X_i \tag{5}$$

where X_i is the total volume that belongs to group G_i

Effort Hours (EH_{ij}) of an agent j on a group i, therefore, becomes a function of the total effective work hours spent in a month by an agent j and the volume handled by him/her in that month

$$\text{EH}_{ij} = \text{TSM}_i / V_{ij} \tag{6}$$

In other words, Effort Hours is the time taken by an agent to resolve one ticket belonging to a specific group. An agent who has lower Effort Hours takes lesser time to complete an incident ticket belonging to the *i*th group.

Agent success rate [ASR] of an agent is a measure of how effective he/she has been in meeting the overall SLA targets. An agent who has a higher success rate is more efficient at meeting overall SLA targets.

$\text{ASR}_{ij} = \frac{\text{Tickets that have met target SLA}}{V_{ij}}$ V_{ij} is the volume handled by an agent j of tickets that belong to group G_i. If Z_i is the number of tickets that have met SLA, then

$$\text{ASR}_{ij} = Z_i / V_{ij} \tag{7}$$

Agent score, which is a measure of the efficiency of an agent in resolve an incident ticket type, can be a function of the Effort Hours (EH_{ij}) & agent success rate (ASR_{ij}) defined as follows:

$$S_j = F\left\{\text{Max}\left(\text{ASR}_{ij}\right), \text{Min}\left(\text{EH}_{ij}\right)\right\} \tag{8}$$

Table 4 shows how we can calculate agent score 4 and Table 3 lists the assumptions used.

Table 3 Assumptions

T_i	100			
Y_i	90%	EWH_i	7	Hrs
p	22	TSD_i	5	Hrs
q	9	TSM_i	114	Hrs
U_i	80%			

Table 4 Agent score calculation

Agents	V_{ij}	Eh_{ij}	Z_i	ASR_{ij}	Agent score (S_j)	Agent rank
Agent 1	30	3.80	10	33	1.3	1
Agent 2	20	5.70	5	25	1.4	2
Agent 3	10	11.40	2	20	2.3	4
Agent 4	5	22.81	5	100	22.8	5
Agent 5	25	4.56	12	48	2.2	3

Additional parameters such as priority and agent tenure may be introduced to further fine-tune the model in the scenarios mentioned.

Scenario 1: In this case, we can use the method of weighted scoring to determine who would be the best agent to work on a Priority 1 ticket. Table 5 shows an example of agent ranking, based on three parameters agent efficiency score (as discussed before), agent tenure, and ticket priority. In this case, Agent 1 has a tenure of 6 months and is relatively new in the system; hence, he is ranked lowest, i.e., 5, whereas Agent 3 has tenure of 3+years and is ranked the highest, i.e., rank 1.

The case considered is for medium priority tickets, i.e., Priority 3, and according to this, if agents have worked on a large volume of Priority 3 tickets and has more experience in solving them, the agent gets a higher rank.

The weights assigned to different parameters on a scale of 1–10 are as follows:

- Agent efficiency weight (w_1): 5
- Agent tenure weight (w_2): 1
- Ticket priority weight (w_3): 4.

Based on the weights, we calculate the weighted total score of all the 5 agents, and the agent with the highest score will be selected/assigned the ticket as explained

Table 5 Agent ranking based on agent efficiency, agent tenure, and ticket priority

	Agent 1	Agent 5	Agent 2	Agent 3	Agent 4
Agent rank	1	2	4	5	3
Tenure rank	5	4	2	1	3
Priority rank	5	1	2	3	4

Table 6 Ticket allocation with agent score

	Agent 1	Agent 2	Agent 3	Agent 4	Agent 5
Score 1 (Agent efficiency score)	40	30	10	5	15
Score 2 (Tenure)	5	10	30	40	15
Score 3 (Priority)	5	40	30	15	10
Total score 1 (TS1)	200	150	50	25	75
Total score 2 (TS2)	5	10	30	40	15
Total score 3 (TS3)	20	160	120	60	40
Total score	225	320	200	125	130
Revised ranking (RSj)	2	1	3	5	4

in Table 6.

$$\mathrm{RS}_j = \sum w_1\mathrm{TS}_1, w_2\mathrm{TS}_2, w_3\mathrm{TS}_3$$

where Sj is score of agent j and RS_j is the ranking of the agent.

Scenario 2: We will also need to consider type and category of an incident while computing the score/ranking of agents.

For e.g., Agent 1 might be working only on password-related issues, while Agent 2 works on Citrix application incidents. Hence, the scoring table needs to be created for different sets of incidents and agents separately. Table 7 provides an example of the same.

In addition to these parameters, the following considerations also need to be taken during ticket assignment.

(a) ***Agent workload***: For the top-ranked agents, we need to calculate the workload of the agents based on a threshold of number of tickets that they will work on, in a day. The workload is calculated as a function of total tickets that an agent is working on.

Total tickets currently allocated to agent = all work in progress (WIP) tickets. Consider Table 8 with ticket details assigned to Agent 2, based on the priority and corresponding daily threshold values.

In this case, since Agent 2 has already crossed the threshold of 7 for Priority 3 tickets in a day, though he is the best agent, he cannot be assigned the new ticket.

Table 7 Agent scoring by type and category of incidents

	Agent 1	Agent 2	Agent 3	Agent 4	Agent 5
Sub-category 1 (Password)	1	2	3	4	5
Sub-category 2 (Share drive)	5	4	3	2	1
Sub-category 3 (Citrix)	3	1	4	5	2

Table 8 Tickets assigned to Agent 2

Priority	P1	P2	P3	P4	P5
Threshold	2	6	7	4	2
Existing (WIP)	0	2	7	0	0

Table 9 Assumptions and notations for agent scoring

Variable	Notation	Value	Comment/calculation
Working days in a month	p	22	After accounting for holidays
Work hours in a day	q	10	
Utilization target	$U\%$	80%	
Effective work hours in a day	EWH	8	$q * U\%$
% of incident tickets handled	$Y\%$	90%	Assuming same proportion for all agents
Total effective hours spent on incident tickets (per day)	TSD	7.2	$Y\%$ * EWH
Total effective hours spent on incident tickets (per month)	TSM	158.4	TSD * p

The ticket will be assigned to the next best agent according to the ranking in Table 3, i.e., Agent 1.

(a) ***Agent availability***: The agent availability needs to be checked from an online roster, and this is the final step in ticket allocation. For e.g., if Agent 1 is about to end his shift and has only 1 h left, and the SLA for the ticket is 4 h, then the ticket needs to be assigned to next available agent, i.e., Agent 3.

Tables 9 and 10 and Fig. 1 illustrate in detail that how our approach will work on a sample data and final agent score get generated.

For illustration on how the real-time allocation will work, refer to Fig. 2.

5 Practical Implementation

We consider two scenarios, one in which the existing workload of an agent is ignored, and another in which due consideration is given to the existing workload of an agent. In the first scenario Fig. 3, assumptions listed below have been considered (Fig. 4):

Scenario 1: Workload is not considered

Assumptions:

(1) All agents have same tenure/experience
(2) All incoming tickets are of the same priority

Table 10 Sample input data (a snippet only)

Start time	Month	Agent name	Group	Priority	Ticket number	SLA Met(Success=1,Failure = 0)	TRT	Ticket status	Ticket description
9/23/2019 2:14:07	9	Agent 44	Group 1	2	Ticket 2	1	5.92	Closed	Description 2
9/14/2019 4:41:21	9	Agent 21	Group 5	2	Ticket 6	1	5.87	Closed	Description 6
8/19/2019 8:47:29	8	Agent 2	Group 5	2	Ticket 11	1	5.73	Closed	Description 11
9/2/2019 8:57:23	9	Agent 28	Group 5	2	Ticket 12	1	5.63	Closed	Description 12
9/14/2019 1:11:37	9	Agent 3	Group 1	2	Ticket 14	1	5.43	Closed	Description 14
8/4/2019 7:58:24	8	Agent 14	Group 5	2	Ticket 16	1	5.42	Closed	Description 16
9/7/2019 0:09:43	9	Agent 2	Group 5	2	Ticket 17	1	5.29	Closed	Description 17
8/4/2019 6:40:55	8	Agent 28	Group 5	2	Ticket 18	1	5.22	Closed	Description 18
9/26/2019 4:26:01	9	Agent 2	Group 5	2	Ticket 20	1	5.11	Closed	Description 20
10/12/2019 8:35:28	10	Agent 18	Group 5	2	Ticket 27	1	4.84	Closed	Description 27

(continued)

Table 10 (continued)

Start time	Month	Agent name	Group	Priority	Ticket number	SLA Met(Success=1,Failure = 0)	TRT	Ticket status	Ticket description
10/10/2019 21:18:36	10	Agent 28	Group 4	2	Ticket 29	1	4.79	Closed	Description 29
9/28/2019 10:21:51	9	Agent 37	Group 5	2	Ticket 30	1	4.78	Closed	Description 30
8/11/2019 12:29:42	8	Agent 30	Group 5	2	Ticket 35	1	4.67	Closed	Description 35
9/19/2019 11:50:22	9	Agent 13	Group 1	1	Ticket 368	0	1.5	Closed	Description 368
10/7/2019 7:39:19	10	Agent 41	Group 5	2	Ticket 369	0	1.5	Closed	Description 369
8/24/2019 7:20:25	8	Agent 33	Group 5	2	Ticket 372	0	1.48	Closed	Description 372
9/21/2019 8:16:15	9	Agent 2	Group 5	2	Ticket 376	0	1.46	Closed	Description 376
10/12/2019 8:12:36	10	Agent 4	Group 5	2	Ticket 377	0	1.46	Closed	Description 377
8/17/2019 4:23:53	8	Agent 42	Group 5	2	Ticket 378	0	1.46	Closed	Description 378
10/5/2019 8:29:29	10	Agent 12	Group 11	2	Ticket 380	0	1.45	Closed	Description 380

(continued)

Table 10 (continued)

Start time	Month	Agent name	Group	Priority	Ticket number	SLA Met(Success=1,Failure = 0)	TRT	Ticket status	Ticket description
9/8/2019 5:17:39	9	Agent 44	Group 1	2	Ticket 381	0	1.44	Closed	Description 381
9/1/2019 8:27:16	9	Agent 15	Group 5	2	Ticket 382	0	1.43	Closed	Description 382
10/6/2019 13:42:13	10	Agent 37	Group 5	2	Ticket 386	0	1.41	Closed	Description 386
10/11/2019 9:53:25	10	Agent 5	Group 5	2	Ticket 387	0	1.4	Closed	Description 387
8/30/2019 9:22:40	8	Agent 13	Group 1	2	Ticket 388	1	1.4	Closed	Description 388
8/11/2019 13:15:23	8	Agent 10	Group 5	2	Ticket 391	0	1.38	Closed	Description 391
9/2/2019 9:20:32	9	Agent 33	Group 5	2	Ticket 398	0	1.33	Closed	Description 398
8/2/2019 2:41:25	8	Agent 21	Group 5	2	Ticket 400	0	1.32	Closed	Description 400
9/8/2019 9:11:07	9	Agent 33	Group 5	2	Ticket 406	1	1.31	Closed	Description 406
8/25/2019 9:56:03	8	Agent 29	Group 5	2	Ticket 689	1	4.18	Closed	Description 689
9/16/2019 9:08:31	9	Agent 22	Group 5	2	Ticket 690	1	4.18	Closed	Description 690

(continued)

Table 10 (continued)

Start time	Month	Agent name	Group	Priority	Ticket number	SLA Met(Success=1,Failure = 0)	TRT	Ticket status	Ticket description
9/15/2019 8:29:08	9	Agent 39	Group 5	2	Ticket 693	1	4.18	Closed	Description 693
9/2/2019 12:36:14	9	Agent 15	Group 5	2	Ticket 695	0	4.17	Closed	Description 695
9/6/2019 6:44:24	9	Agent 27	Group 5	2	Ticket 697	0	4.17	Closed	Description 697
8/9/2019 10:09:10	8	Agent 22	Group 5	2	icket 699	0	4.16	Closed	Description 699

Following Calculations are only done for Group 5 and takes months 8,9, 10 into consideration

EFFORT HOURS (EH)

Agent Name	Volume (Month 8)	Volume (Month 9)	Volume (Month 10)	Avg Vol(V)	Effort Hours(EH =TS M/V)	Comments
Agent 10	10			10	16	Not considering since agent is not present month 9,10
Agent 11			4	4	40	
Agent 12		7	8	8	21	
Agent 14	15	20		18	9	Not considering since agent is not present month 10
Agent 15	19	16	10	15	11	
Agent 18	12	30	16	19	8	
Agent 19			8	8	20	
Agent 2	23	30	6	20	8	

Following Calculations are only done for Group 5 and takes months 8,9, 10 into consideration

AGENT SUCCESS RATES (ASR)

Agent Name	Total Success	Total Tickets	Agent Success Rate(ASR)
Agent 10	6	10	60%
Agent 11	2	4	50%
Agent 12	9	15	60%
Agent 14	22	35	63%
Agent 15	21	45	47%
Agent 18	30	58	52%
Agent 19	2	8	25%
Agent 2	35	59	59%

Agent Score is calculated as function of ASR and EH Here we assign relatives ranks to agents based on EH & ASR. Those with lower EH gets higher EH Score. Those with higher ASR gets higher ASR Rank.

AGENT SCORE (S) [Assuming equal weightage for EH Score and ASR Score]

Agent Name	EH Score	ASR Score	Agent Score
Agent 10	37	43	40
Agent 11	26	32	29
Agent 12	29	43	36
Agent 14	47	44	45.5
Agent 15	44	28	36
Agent 18	48	34	41
Agent 19	31	23	27
Agent 2	50	41	45.5

Final Agent List

Agent Rank [Best Agent gets rank 1]

Agent Name	Rank
Agent 2	1
Agent 25	2
Agent 27	2
Agent 28	4
Agent 18	5
Agent 29	5
Agent 39	7
Agent 22	8

Fig. 1 Calculation of agent score and final agent list

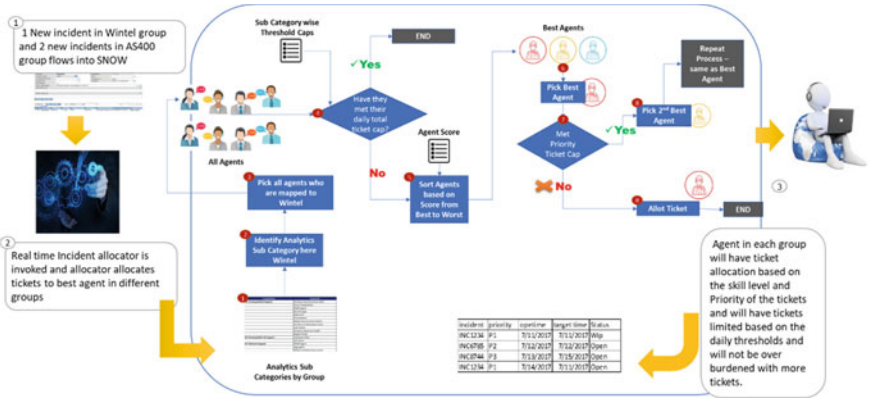


Fig. 2 Real-time agent allocation by agent scoring

There is no limit/cap on how many tickets an agent can resolve in a day. In this scenario 1, the ticket type 1, gets assigned to Agent 2, since he has the least Effort Hours for type 1 tickets, among all the agents present. If a second ticket had come to him of the same type, it would still be assigned to Agent 2. Similarly, the type 2 ticket also gets assigned to Agent 2, since he has the least effort in time units to resolve the ticket. To re-iterate, we only consider the Effort Hours, since we are assuming all agents have the same success rate. This method of allocation will have the obvious flaw, that a single agent might end up being over-loaded with tickets, thereby increasing the overall resolution time. This will be counterproductive to achieving efficiency in the allocation and resolution of tickets.

Scenario 2: Agents are one ticket away from reaching their workload thresholds

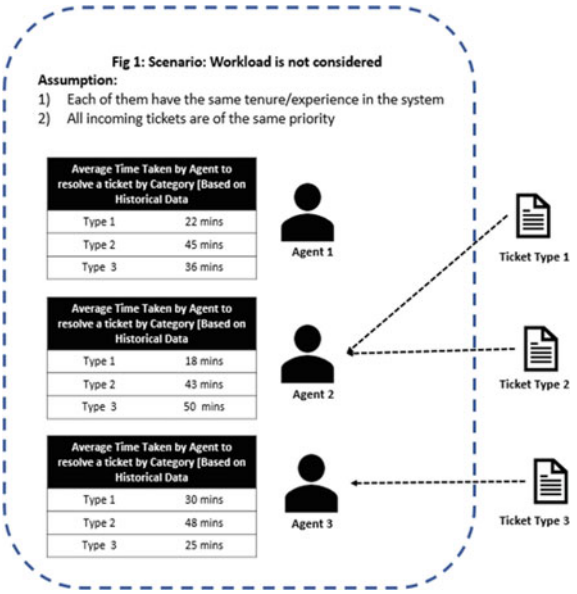


Fig. 3 Scenario 1 (workload is not considered)

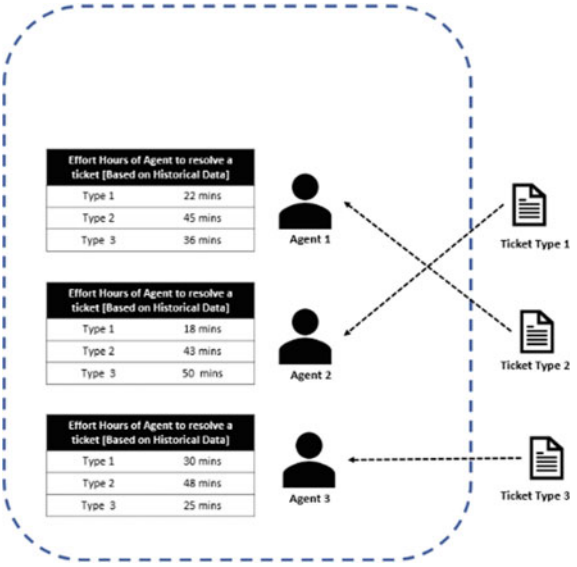


Fig. 4 Scenario 2 (agents are one ticket away from reaching their workload thresholds)

Assumptions:

(1) All agents have same tenure/experience
(2) All incoming tickets are of the same priority

To prevent burdening few agents with all the tickets, even if they are the best agents to resolve them, we introduce thresholds on the workload, an agent will have, at any point of time. For ease of exposition, we will assume that all agents are one ticket away from reaching their workload thresholds. In this scenario, the ticket of type 2 will flow to agent 1, since he has the second least effort in terms of time unit to resolve the ticket. Capping the workload of an agent will ensure that no agent is overworked. Further, to make the allocation more robust, we can introduce other parameters such as priority of the ticket and agent tenure into play; this can be in the form of workload thresholds by priority or by tenure, or by giving weightage to each of the factors in the total agent score as discussed in Sect. 4—Our Approach.

6 Conclusion

Ticket handling time can be substantially reduced by minimizing human intervention and bias in the allocation decision. While in the as-is process, a team lead/queue manager allocates a ticket to an agent, taking into consideration the workload, often the perceived ability of the agent is subject to bias. Allowing agents to pick up tickets directly from the common queue eliminates the bias on perceived ability but introduces externalities arising from cherry-picking. Agents might resort to picking up those tickets that require lesser time to complete and/or effort with an explicit intent to keep their average resolution time low and/or customer satisfaction rates and success rates high. Adopting a data-driven approach gives two-pronged benefits, first in reducing the time for allocation, and second in reducing the resolution time, by utilizing the edge, an agent might have in resolving a ticket. By using the output from the agent scorer, the "incident allocation engine" eliminates the need to have a queue manager. In our current version of the solution, we are targeting "first time allocation." For our future work, we can further enhance the solution by including incident re-allocation by introducing "incident re-assignment rules." For e.g., incident sitting in a given agent's queue can be reassigned to another agent based on workload and thus bring in ongoing real-time optimization. For e.g., a ticket P5 was assigned to Agent 1 initially based on the rules. Now, the agent 1 has 3 P2 tickets, while Agent 2 has only 3 P5 tickets. The P5 ticket assigned to Agent 1 can then be moved to Agent 2, so that Agent 1 can focus on the high-priority tickets. This may mean re-looking at the agent queues to do near real-time optimization.

References

1. Zhou, W., Tang, L., Zeng, C., Li, T., Shwartz, L., Grabarnik, G. Y. (2016) Resolution recommendation for event tickets in service management. *IEEE Transactions on Network and Service Management, 13*(4), 954–967. https://doi.org/10.1109/tnsm.2016.2587807.

2. Deshpande, P. M., Garg, D., Suri, R. M. (2008). Auction based models for ticket allocation problem in it service delivery industry. In *IEEE International Conference on Services Computing*, USA, July 2008. https://doi.org/10.1109/scc.2008.83. ISBN: 978-0-7695-3283-7.
3. Graham, R. L. (1966). Bounds for certain multiprocessing anomalies. *The Bell System Technical Journal, 45*(9), 1563–1581. https://doi.org/10.1002/j.1538-7305.1966.tb01709.x.
4. Bartal, Y., Fiat, A., Karloff, H., & Vohra, R. (1995). New algorithms for an ancient scheduling problem. *Journal of Computer and System Sciences, 51*(3), 359–366.
5. Karger, D., Philips, S., Torng, E. (1994). A better algorithm for an ancient scheduling problem. In *ACM DL, SODA*, USA, 1994 (pp. 132–140). ISBN: 0-89871-329-3.
6. Albers, S. (1999). Better bounds for online scheduling. In *29th Annual ACM Symposium on Theory of Computing, 1997*, Vol 29, no 2, Oct. 1999, pp. 459–473.
7. Lim, K., Kangbok, L., & Soo, C. (2011). Improved bounds for online scheduling with eligibility constraints. *Theoritical Computer Science, 412*(2011), 5211–5224. https://doi.org/10.1016/j.tcs.2011.05.029.
8. Li, R., & Huei-Chuen, H. (2007). On-line scheduling on identical machines for jobs with arbitrary release times. In *Research Gate*, Dec 2007. https://doi.org/10.5772/5217.
9. Bartolini, C., Stefanelli, C., & Tortonesi, M. (2008). SYMIAN: A simulation tool for the optimization of the IT incident management process. In *DSOM* (pp. 83–94). https://doi.org/10.1007/978-3-540-87353-2_7.
10. Bhandari, A., Scheller Wolf, A., & Harchol-Balter, M. (2008). An exact and efficient algorithm for the constrained dynamic operator staffing problem for call centers. In *ACM DL, Journal Management Science, 54*, USA (pp. 339–353). https://doi.org/10.1287/mnsc.1070.0819.

An Efficient Makespan Reducing Task Scheduling Algorithm in Cloud Computing Environment

Dhananjay Kr. Sharma, Dhirendra Kr. Shukla, Vijay Kr. Dwivedi, Avadhesh Kumar Gupta, and Munesh C. Trivedi

Abstract The current trends in task scheduling problems in cloud computing are moving toward the optimization of task execution time with the invention of novel approaches for heterogeneous environments. This article aims to decrease the makespan time of the scheduling in cloud computing environment. The article introduces an approach for efficient task scheduling of the diversified machines used in cloud to minimize the makespan time. The proposed algorithm was checked with the Braun benchmark dataset. The experimental results demonstrate that the proposed algorithm minimizes the overall makespan up to 11.87% as compared to the other recent implemented algorithms.

Keywords Cloud computing · Task scheduling · ETC · ECTC · MaxUtil · ETSA

1 Introduction

Cloud computing is a recent trend which provides different services on pay-per-use basis. There is a lot area where cloud computing is playing major roles. Cloud scheduling is one of the major challenges nowadays [1, 2]. There are various kinds of cloud services like infrastructure as service (IaaS), platform as service (PaaS), software as service (SaaS), etc. In IaaS, users use the cloud services via virtual

D. Kr. Sharma (✉)
United Institute of Technology, Prayagraj Uttar Pradesh, 211009, India
e-mail: dhananjay7862@gmail.com

D. Kr. Shukla
MNNIT Allahabad, Prayagraj Uttar Pradesh, 211004, India

V. Kr. Dwivedi
United College of Engineering and Research, Prayagraj Uttar Pradesh, 211009, India

A. K. Gupta
School of Computing Science and Engineering, Galgotias University, Greater Noida, India

M. C. Trivedi
National Institute of Technology, Agartala 799046, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_31

machines (VMs) deployed in their data centers [3, 4]. Cloud computing provides an excellent growth in small-to-medium and start-up businesses because it provides the limit-free services. The demand of cloud computing rapidly grows day by day. So that, the management of cloud resources is an important task so that we can utilize them as much as possible. The advantages of cloud computing are that it solves the problems without having any kind of physical infrastructure.

The actual implementation of cloud is a big challenge because it requires a lot of resources like high configuration machines, a huge number of storage space, and many more. So, we require a tool/framework which provides the facility for implementing the cloud and the solution; this problem is CloudSim or Python which provides a lot number of classes and methods by which we can implement the cloud computing systems.

2 Related Work

There are many works been proposed in recent years for task scheduling. Panda et al. [5] presented a multi-objective task scheduling algorithm. They have been shown to require $O(kl^2m)$ time for k iterations. Singh et al. [3] focused on issues and challenges of the resource scheduling. They have proposed cloud-based task scheduling algorithms. Chun-Wei et al. [6] also presented a survey on the meta-heuristic scheduling for cloud. In summary, they conclude that the definition of cloud is still not clear and the proposed algorithm is just hypothesis. When cloud will more mature then the proper algorithm of cloud could design. Shukla et al. [7] presented algorithm for task scheduling to minimize the makespan and resource utilization. Panda et al. [8] presented a new heterogeneous multicloud environment algorithm. The algorithm showed requiring O(logn) time for one task and n VMs. They extended their work in [1] and presented an efficient task consolidation algorithm for cloud computing systems where they presented a task consolidation algorithm multi-criteria-based task consolidation (MTC) for cloud computing systems. The algorithm has been shown to require *O(knmp)* time. The authors of [2] also present an efficient allocation of resources algorithm (ERAA) for IaaS cloud computing systems. Here, the experimental results have been compared with two well-known existing resource allocation algorithms (HAIZEA [9] and Nathani [10]).

Liu et al. [11] proposed algorithm for mobile cloud environments, and the algorithm name was deadline-constrained multi-objective task scheduling. In this paper, he proposed algorithms (HEFT-T) for the unconstrained and deadline-constrained task scheduling in mobile cloud computing environments. Sotiriadis et al. [12] proposed a new VM placement algorithm based on the previous virtual machine experiences. They present an algorithm for self-management of virtual machine placement based on utilization levels of physical machine (PM) and virtual machine (VM). Rani et al. [13] presented the basic analysis on CloudSim simulator and also have done some simple examples for the experiment purposes. Humane et al. [14] conclude that the CloudSim not only could be used for basic datacenters but also

could be used in research for simulation of research problems. Pratap et al. [15] implemented FCFS, round-robin, and SJF task scheduling algorithms for the cloud using CloudSim and compared the analytical results.

There are multiple companies (such as Amazon and Netflix) providing cloud resources to the customers. Day by day, the cloud users are increasing with a rapid rate. So, in this study we will observe about the best algorithm which could be used for cloud scheduling.

The purpose of this research is to learning and comparing the cloud scheduling algorithms. In this paper, we have just implemented different cloud scheduling algorithms like first come first serve (FCFS), round-robin, particle swarm optimization (PSO), and shortest job first (SJF) algorithms. Here, we will observe which algorithm performance is better than the others. For implementing the algorithm, we will use the CloudSim tool with the NetBeans IDE.

3 Proposed Algorithm

The proposed algorithm aims to enhance the makespan time of different machines. The proposed algorithm uses the makespan time as the parameter to evaluate the performance of several algorithms.

- Makespan: It is defined as the maximum time at which a machine finishes its computation of all the assigned tasks allocated to that machine with the assumption that all nodes start execution of computation at the same time.

$$\text{Makespan} = \sum_{i=1}^{m} \text{ETE}(i, j)$$

where m is total number of machines, and i and j represent the corresponding task with execution time.

The proposed algorithm along with energy-aware task consolidation (ETC) [16], energy-conscious task consolidation (ECTC) [16], MaxUtil [17], and energy-efficient task scheduling algorithm (ETSA) [18] has been applied on the Braun et al. [19] dataset to observe the makespan time of various machines. The results obtained from the implemented experiment have been compared with results generated by the implementation of the authors Panda et al. [18]. The proposed algorithm's flowchart is as follows (Fig. 1).

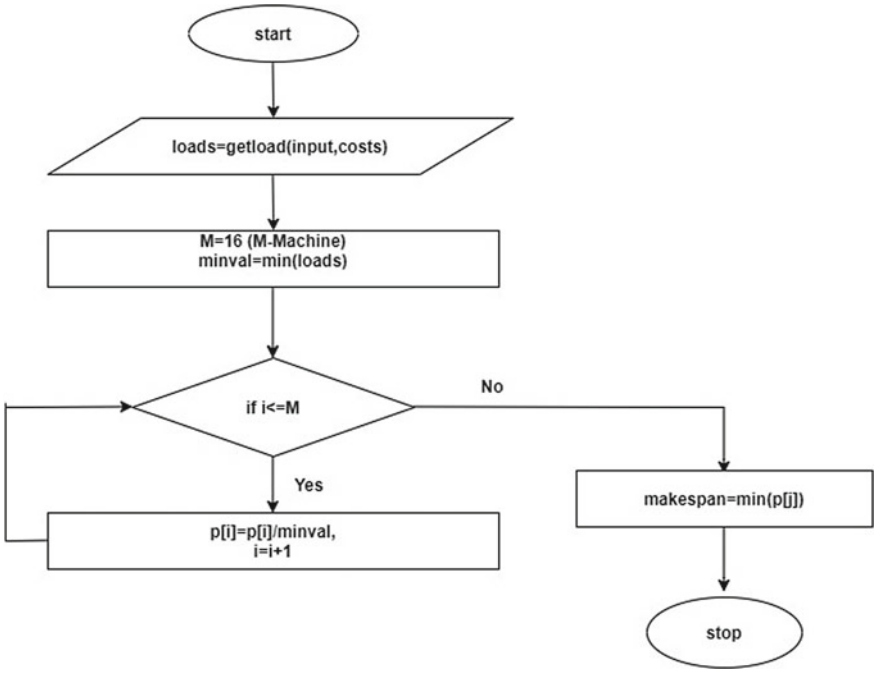


Fig. 1 Proposed algorithm

4 Experimental Results

The proposed algorithm applied on Braun dataset minimizes makespan time as compared to the other implemented algorithms in this experimental setup. Table 1 shows the comparison of makespan time of all the implemented algorithms.

Figure 2 illustrates the consistent machines with their makespan time. The chart shows that the ETSA is better among all other implemented algorithms for four machines implemented in this experimental setup. The diagram also highlights that the proposed algorithm reduces the makespan time by significant percentages like 9.79%, 13.37%, 10.81%, 9.52% for the implemented four machines, respectively.

Figure 3 illustrates the comparative analysis of the inconsistent machines with their makespan. It shows that the ETSA is better among all other implemented algorithms for the four implemented machines in this experimental setup. The chart also reveals that the proposed algorithm reduces the makespan time for inconsistent machines by significant amount of percentages like 13.58%, 14.85%, 15.38%, 10.53% for the implemented four machines, respectively.

Figure 4 provides the comparative analysis of the semi-consistent machines with their makespan time. It highlights that ETSA algorithm performs better than all other implemented algorithms for the currently implemented four machines in this experimental setup. The chart emphasizes that the proposed algorithm minimizes the

Table 1 Comparison of makespan (in seconds) of ETC, ECTC, MaxUtil, and ETSA with proposed algorithm

Machines	ETC	ECTC	MaxUtil	ETSA	Proposed algorithm	Percentage difference (ETSA-proposed algorithm)
m_c_hihi	1.12E + 06	1.46E + 06	4.98E + 05	1.18E + 05	1.07E + 05	9.79
m_c_hilo	1.00E + 05	1.49E + 05	4.50E + 04	1.17E + 04	1.03E + 04	13.37
m_c_lohi	1.02E + 03	1.31E + 02	9.91E + 02	3.90E + 01	3.50E + 01	10.81
m_c_lolo	1.02E + 03	8.66E + 02	9.80E + 02	1.10E + 01	1.00E + 01	9.52
m_i_hihi	9.03E + 04	7.21E + 05	1.50E + 05	2.75E + 04	2.40E + 04	13.58
m_i_hilo	7.29E + 05	8.50E + 04	4.51E + 04	2.92E + 03	2.51E + 03	14.85
m_i_lohi	1.54E + 02	8.80E + 01	4.90E + 01	7.00E + 00	6.00E + 01	15.38
m_i_lolo	4.74E + 02	3.70E + 02	1.80E + 01	2.00E + 00	1.80E + 02	10.53
m_s_hihi	8.47E + 05	1.52E + 06	5.00E + 05	8.48E + 04	7.92E + 04	6.88
m_s_hilo	4.45E + 05	1.83E + 05	7.35E + 04	8.19E + 03	7.53E + 03	8.43
m_s_lohi	7.03E + 02	1.70E + 02	4.08E + 02	2.30E + 01	2.00E + 01	13.95
m_s_lolo	5.12E + 02	6.11E + 02	4.09E + 02	7.00E + 00	6.00E + 03	15.38

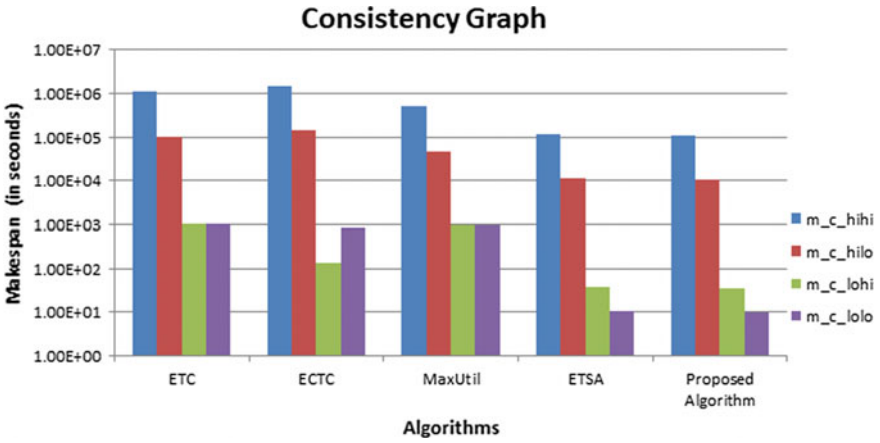


Fig. 2 Graphical representation of the makespan for consistent machine

makespan time for semi-consistent machines by significant amount of percentages like 6.88%, 8.43%, 13.95%, 15.38% for the implemented four machines, respectively.

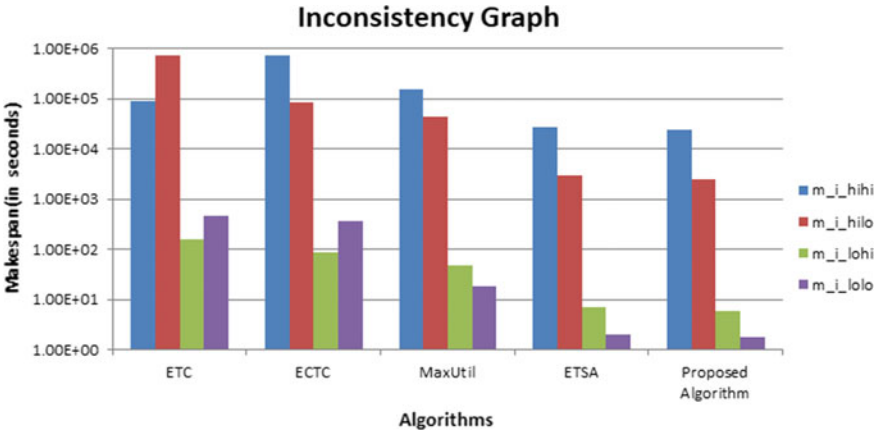


Fig. 3 Graphical representation of the makespan for inconsistent machine

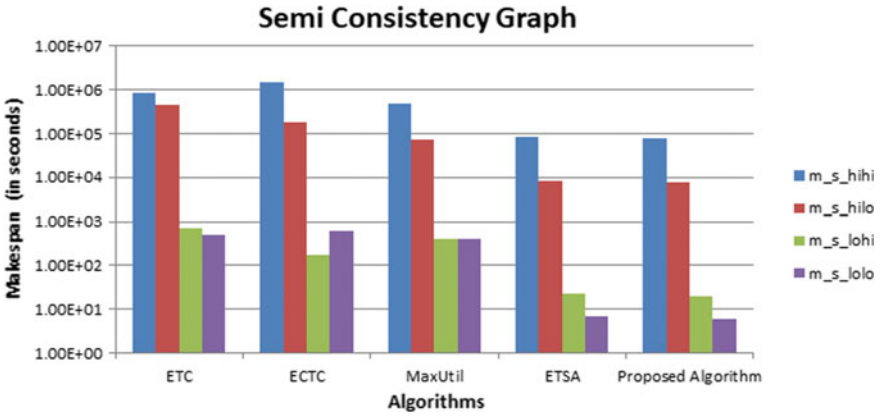


Fig. 4 Graphical representation of the makespan for semi-consistent machine

5 Conclusion

This article proposes an efficient task scheduling algorithm in heterogeneous cloud computing environments. The algorithm minimizes the overall makespan for the twelve machines employed in the implemented experimental setup. The analysis of experimental results highlights that the proposed algorithm significantly decreases the makespan for all used twelve machines on Braun dataset.

References

1. Panda, S. K., Jana, P. K. (2016). An efficient task consolidation algorithm for cloud computing systems. In N. Bjørner et al. (Ed), Springer International Publishing Switzerland 2016: ICDCIT 2016, LNCS 9581.
2. Panda, S. K., & Jana, P. K. (2015). An efficient resource allocation algorithm for IaaS cloud, Springer International Publishing Switzerland 2015. In R. Natarajan et al. (Ed), ICDCIT 2015, LNCS 8956, pp. 351–355.
3. Singh, S., & Chana, I. (2016). A survey on resource scheduling in cloud computing: Issues and challenges. *Journal of grid computing, 14,* 217–264. https://doi.org/10.1007/s10723-015-9359-2.
4. Panda, S.K., & Jana, P. K. (2016). Normalization-based task scheduling algorithms for heterogeneous multi-cloud environment. New York: Springer Science+Business Media.
5. Panda, S. K., Jana, P. K., A multi-objective task scheduling algorithm for heterogeneous multi-cloud environment. https://doi.org/10.1109/edcav.2015.7060544.
6. Tsai, C. -W., & Joel, J. P. C. (2014). Metaheuristic scheduling for cloud: A survey. *IEEE Systems Journal, 8*(1).
7. Shukla D. K., Kumar, D., & Kushwaha, D. S. An efficient tasks scheduling algorithm for batch processing heterogeneous cloud environment. *International Journal of Advanced Intelligence Paradigms*. [Inderscience (Scopus Indexed)] https://doi.org/10.1504/ijaip.2021.10027089.
8. Pandaa, S. K., Gupta, I., & Jana, P. K., Allocation-Aware task scheduling for heterogeneous multi-cloud systems. In *2nd International Symposium on Big Data and Cloud Computing (ISBCC'15).*
9. Haizea, http://haizea.cs.uchicago.edu/whatis.html. Accessed on 9 Jan 2014.
10. Nathani, A., Chaudhary, S., & Somani, G. (2012). Policy based resource allocation in IaaS cloud. *Future Generation Computer Systems, 28,* 94–103.
11. Liu, L., Fan Q., & Buyya, R. (2018). A deadline-constrained multi-objective task scheduling algorithm in mobile cloud environments. (pp. 2169–3536). IEEE.
12. Sotiriadis, S., Bessis, N., & Buyya, R., Self managed virtual machine scheduling in Cloud systems, S0020-0255(17)30827-7.
13. Ekta Rani, Harpreet Kaur, Study on fundamental usage of cloudsim simulator and algorithms of resource allocation in cloud computing. (p. 40222). IEEE.
14. Humane, P., & Varshapriya, J. N. (2015). Simulation of cloud infrastructure using cloudsim simulator: A practical approach for researchers. 978-1-4799-9855-5/15/$31.00 ©2015 IEEE.
15. Pratap, R., Zaidi, T., Comparative study of task scheduling algorithms through cloudsim. 978-1-5386-4692-2/18/$31.00 ©2018 IEEE.
16. Hsu, C. H., Slagter, K. D., Chen, S. C., & Chung, Y. C. (2014). Optimizing energy consumption with task consolidation in clouds. *Information Sciences, 258,* 452–462.
17. Lee, Y. C., & Zomaya, A. Y. (2012). Energy efficient utilization of resources in cloud computing systems. *Journal Supercomput, 60,* 268–280. https://doi.org/10.1007/s11227-010-0421-3.
18. Panda, S. K., Jana, P. K. (2018). An energy-efficient task scheduling algorithm for heterogeneous cloud computing systems. Received: 5 April 2018/ Revised: 29 August 2018/Accepted: 17 October 2018/Published online: 30 October 2018, Springer Science+Business Media, LLC, part of Springer Nature 2018.
19. Braun, F.N. (2018). https://code.google.com/p/hcsp-chc/source/browse/trunk/AE/ProblemInstances/HCSP/Braun_et_al/u_c_hihi.0?r=93 Accessed on 9 Mar 2018.
20. Panda, S. K., Gupta, I., Jana, P. K. (2017). Task scheduling algorithms for multi-cloud systems: Allocation-aware approach. New York: Springer Science+Business Media.

Brain-Related Diseases and Role of Electroencephalography (EEG) in Diagnosing Brain Disorders

Puja Dhar and Vijay Kumar Garg

Abstract Electroencephalography (EEG), one of the most effective method for analyzing the electrical activities of brain. It is used widely in diagnosis of brain-related diseases. There are many disorders of brain which are discussed in this paper. The datasets and the number of subjects which were included in these were also reviewed. Review of various research papers and articles published in refereed peer-reviewed international journals in the domain of neurological science is also taken into consideration. A review has been developed that emphasized on the neurological disorders and role of electroencephalography in their detection. Role of electroencephalography is also studied for diagnosing the disorders of brain.

Keywords EEG · Epilepsy · Seizure · Alzheimer · Classification · Disorders · Neurological · Datasets

1 Introduction

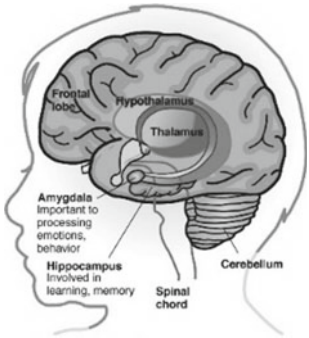
Brain is a considered as network of complex network of neurons. To know about the diseases related to the brain, one should have knowledge of working and structure of brain, and how it responds either adaptively or maladaptively. To know about these disorders, the analysis of the electroencephalogram (EEG) is one of the solutions to properly diagnose these disorders. There are many disorders related to brain including autism, Alzheimer's disease, dementia, seizure, tumor, Scoliosis, etc.

Autism is defined as widespread development disorder of brain as shown in Fig. 1. It is considered widespread because this disease affects several features of cognition and behavior and growth. The autism starts its effect from the birth of a child. To understand the neurodevelopment mechanism in autism genetic study is useful. This disease affects the genes which are helpful in nervous system development.

P. Dhar (✉) · V. K. Garg
Lovely Professional University, Phagwara, Punjab, India
e-mail: pujadhr@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_32

Fig. 1 Autism



Alzheimer's disease (AD) is the most common form of dementia as seen in Fig. 2. Study says that 25–36 million people are affected with Alzheimer worldwide. Environment and genetic factors both are the causes of this disease. This disease shows the symptoms of hypertension, heart disease, depression, diabetes, and sometimes arthritis too. Dementia is a medical disorder which is considered worldwide as a mental damaging. It is related to the damage in functional abilities and sometimes shows psychiatric troubles and behavioral issues (Fig. 3).

When there are not normal electrical activities in the brain, it causes seizures which are occurred primarily in the cortex as shown in Fig. 4. The clinical observation says that it affects the area of cortex where actually the seizures develop. Over past 35 years brain tumors, Fig. 5, have increased and, treatments have not much improved the predictions of deadly brain tumors. Figure 6 showing the scoliosis which is spinal cord related disorder, and it indicates the uneven shoulders, trouble

Fig. 2 Alzheimer's disease

Fig. 3 Brain with dementia

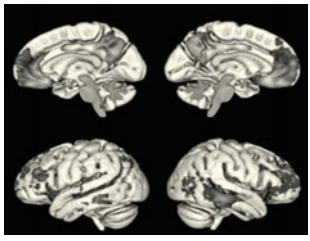


Fig. 4 Seizure disorder

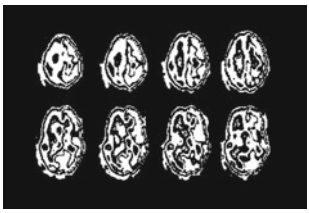


Fig. 5 Brain tumors

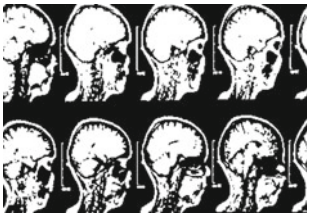
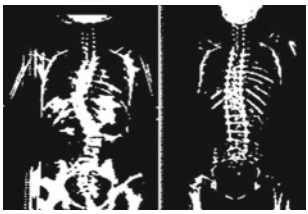


Fig. 6 Scoliosis



walking, shortness of breath, height disorder, numbness, weakness or pain in the legs, and bump in lower back. Spina bifida is a disease which is related to the spinal cord. It is a neural tube problem in which spinal cord is not properly developed. It is related to baby's embryo which finally becomes the brain.

As seen in the Table 1, which shows the rates of seizures depending upon the age of a person revealing that data which is taken from a longitudinal research of 5000 kids [1]; the frequency of seizure in autism is higher throughout childhood. As per the researcher, the consolidated risk for developing seizure in their sample was approximately 29%.

Table 1 Frequency of seizures according to the age

	Age of person (years)					
	<4	4–7	8–11	12–15	16–19	20 onwards
Number of initial seizure	15	8	5	6	4	1
Person years of observation	546.3	581.4	477.1	342.4	223.5	184.5
Yearly frequency per 1000	29.2	15.3	7.5	20.3	17.7	5.3

2 Literature Review

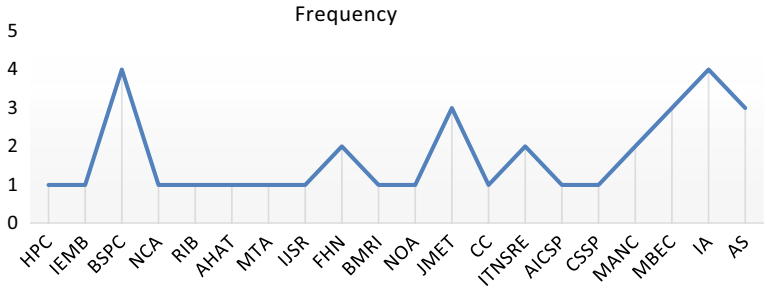
Epilepsy is one of the most common neurological diseases worldwide, and seizure is one of the characteristics of epilepsy. A person having seizure shows the sudden changes in the functioning of brain and shows different behaviors including jerky movements, memory loss, breathlessness, and unconsciousness Siuly et al. [2]. To understand epilepsy, EEG is considered the best tool. List of journals is also mentioned in Table 2 which shows the name of journals and used abbreviations.

As per Qiu et al. [3] , epilepsy affects both physical and mental health of a person. There is an important role of electroencephalogram (EEG) in diagnosing epilepsy. Gotman [4] addressed that half wave signals and peak amplitude, slope, sharpness values are used for the detection of epilepsy. While reviewing the literature Kulkarni [5] addressed that in today's world some of the diseases are expensive, and Alzheimer is considered one of them. EEG is considered as an economical tool but still not giving best results in case of Alzheimer.

Number of studies has been done in regard of epilepsy Holler et al. (2017). The first measure of interaction, coherence has been used to find epilepsy. Kaur et al. [6] have discussed in their paper that EEG signals are helpful to analyze the physiological changes. Researchers have used different emptions and proposed an emption recognition framework using EEG. In this paper, they have used HMV and SVM

Table 2 List of journals with their abbreviation

Name of journal	Abbreviation	Name of journal	Abbreviation
Hindawai publishing corporation	HPC	Bio medical research international	BMRI
IEEE engineering in medicine and biology	IEMB	Neurobiology of aging	NOA
Biomedical signal processing and control	BSPC	Journal of medical engineering and technology	JMET
Neural computing and applications	NCA	Cluster computing	CC
Artificial intelligence and robotics	RIB	IEEE transactions on neural systems and rehabilitation engineering	ITNSRE
Advances in human aspects of transportation	AHAT	Analog integrated circuits and signal processing	AICSP
Multimedia tools and applications	MTA	Circuits system signal process	CSSP
International journal of science and research	IJSR	Multidisciplinary approaches to neural computing	MANC
Frontiers in human neuroscience	FHN	Medical and biological engineering and computing	MBEC
Applied Sciences	AS	IEEE Access	IA



Graph 7 Name of journals versus frequency of research papers reviewed

classification. Kumar et al. [7] addressed that 1% of world's population is affected by epilepsy.

Maksimenko et al. addressed that EEG is a best tool to recognize and represent emotions in human beings more accurately. EEG is helpful to know emotional activities which provoke the brain to produce signals Brunner et al. [8]. Bao et al. [9] mentioned that developing countries are mostly affected by epilepsy, and approximately 85% people are victim of it [10].

3 Brain Disorders

As we know that our body is controlled by brain, which is the part of nervous system, and this system includes the nerves, neurons, and also the spinal cord. Nervous system is responsible for controlling our body senses. There are many brain disorders which can affect our whole body. Fig. 8 shows different brain related disorders and other diseases related to these disorders.

4 Datasets

Every research requires the data on which analysis of measurement of data is done. Datasets are helpful which combination of parameters and models is. Nowadays, the dataset is taken as the primary source of output for the research. Some of the datasets are seriously considered for the study (Table 3).

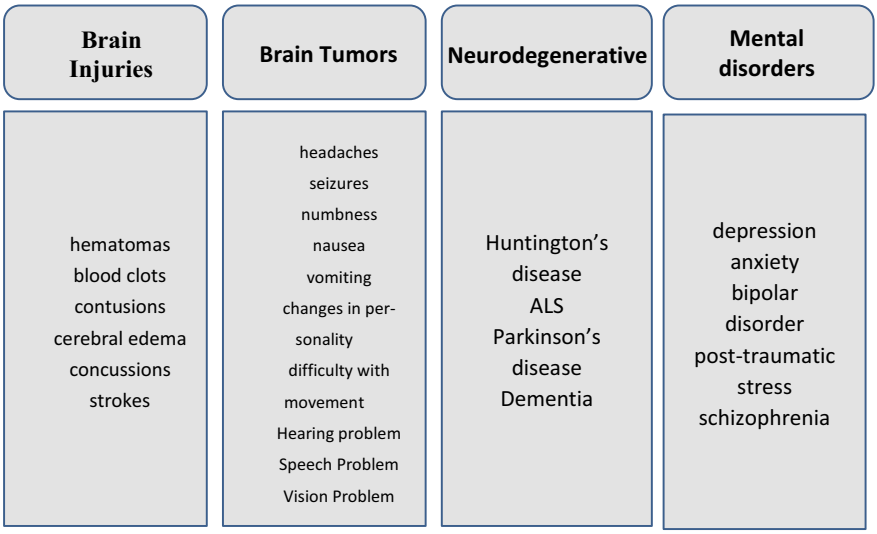


Fig. 8 Some brain related disorders

Table 3 Description of some datasets used by the researchers

Dataset	No. of subjects	Description
Temple university hospital repository	12,000	16-channel EEG EDF files
Psychophysics (700 Mb):	122	64 channel (Alcoholic and controls performing a visual matching task). MATLAB scripts to import to EEGLAB are available here
EEG dataset published on PhysioNet	109	From Gerwin Schalk's team at the Wadworth center in Albany, NY
DEAP dataset	32	EEG (and other modalities) emotion recognition
Continuous EEG	100	Few seconds of 64-channel EEG recording from an alcoholic patient
The DREAMS sleep spindles database	8	A collection of file and sleep scoring only 5 EEG channels
EEG database for BCI applications	10	Various experiments are featured
Motor imagery data	20	Motor imagery data for BCI project (MATLAB files)
LEMON dataset	277	Particularly suited to comprehensively relate cognitive and emotional traits or states to physiological characteristics of brain and body
EEG database dataset	122	Study to examine EEG correlates of genetic predisposition to alcoholism

5 Methods

King et al. (1998) have used clinical data of 300 adults and children for EEG test. The duration was 24 h of the seizure, where it was diagnosed that about 232 patients have epilepsy syndrome. It is also analyzed that 24 h duration is best time duration for the diagnosis. 16-channel analog or 24-channel digital recordings with bipolar longitudinal and traverse montages, standard 10–20 electrode placements were used in their research. While discussing the uses and abuses of EEG, Fowel et al. [11] reviewed the evidences of site searches on EEG, electroencephalography. Guo et al. [12] used down sampled to 128 Hz sample rate. A band pass filter between 4 and 45 Hz is used to remove noise. They focused on valence–arousal plane emotion states; the frequency bands chosen were alpha, beta, and gamma bands. SVM classifier is used with three combinations of the extracted features. It is reviewed that EEG signals have intra-channels dependencies. Differential pulse code modulation (DPCM) method is applied to decrease the redundancy of channel, and appropriate data is extracted. The range of this data is smaller than original one. Mumtaz et al. [13] presented two groups of participants recruited from the outpatients of HUSM, Malaysia. In this study, feature selection remains a challenging research topic and carries critical importance during data analysis involving a typical machine learning (ML) methodology. Jisu et al. [14] addressed brain as a chaotic dynamic system as a nonlinear dynamic system and also discussed about evolution of chaos theory. This paper discusses the relevance of chaotic analysis of EEG signals. EEG is the most important monitoring methodology of epileptic seizure diseases. In this paper, database of Bonn University and Bern-Barcelone EEG is taken. Some researchers proposed technique using three steps including decomposition, feature extraction, and classification. It is also reviwed that use of pre-processing in eliminating the unwanted noise while identifying the seizure is possible. Bonn University EEG database is a noiseless database. Researchers have used VMD approach for decomposition of quasi-orthogonal signal.

6 Classification

Many researchers have used different methods of classifications including sensitivity, specificity, classification accuracy and using these work structure best results are derived using one input layer with five neurons equals to the number of line length features extracted for each EEG segment using MATLAB software. Many classification techniques are used by the researchers including SVM with nonlinear kernel and nave Bayes [1], linear discriminate analysis, logistic regression, KNN.

During study, it is found that there is shortage of studies on the effectiveness of devices used for detection of epilepsy [1]. Many new technologies have been used to detect epilepsy but they are hardly verified by EEG. In some studies, numbers are misleading like a study by Narachania et al. [10] has taken test on 79 patients

Table 4 Result of classification accuracy

Result of classification accuracy to improve EEG bands				
	feature			
Classifier	Theta (%)	Beta (%)	Delta (%)	Alpha (%)
LR	69	65	65.5	69.5
LDA	69	71	65.5	72.5
KNN	69	70	70	72.5

but only 18 seizures were recorded., and Arnedo et al. [15] has taken 64 patients in which only 8 were recorded.

To define the classifications, many terms are applied like True positive (TP), False Negative (FN), True Negative (TN), and False Positive (FP). After taking the tests, it is checked that if a person is having this disease, then test is considered as true positive (TP) as mentioned in Table 4, and if patient is not having this diseases, then diagnosis test declares that a person is free from diseases which means true negative (TN), and so on. This paper addressed the work done by Qiu et al. [3]. Researchers have used denoising sparse autoencoder (DSAE) and have imported pre-processed training data to it. Classification was also applied to the top of DSAE, and finally, epilepsy detection result is obtained. In recent years, automatic seizure detection methods have been proposed by Qiu et al. [3].It was initiated by Gotman [16] and after collaboration with Qu [17] automatic seizure detection method was developed. In this paper, a new system was proposed in which raw EEG signals are pre-processed, and after feature computation, feature selection is done, and its input is sent for testing, classification to diagnose the classes of diseases. Jacob et al. [14] have taken the data from 25 patients of Government Medical College, Thiruvanthapuram, Kerala. EEG waveforms were recorded using Nicolet Nic Vue 3.0 software. 12 s duration was fixed for epochs, and 500 Hz sampling frequency was fixed during recording. LPF-TVD filtering is used for pre-processing. MTLAB software is also used to calculate fractal dimensions (Table 5).

Table 5 Specificity, sensitivity, and accuracy

Outcome of diagnosed tests	Diseases as determined by the SOT		
	Positive	Negative	Row total
Positive	TP	FP	TP + FP
Negative	FN	TN	FN + TN
Column total	TP + TN	FP + TN	TP + FP + FN + TN

7 Research Agenda and Issues Addressed

In recent years, most attention with highest frequency on sample papers reviewed has been to detect epileptic seizure using electroencephalogram (EEG) and finding epilepsy characteristics. It is found that some research papers on neurodegenerative disorders like Alzheimer and epileptic form activity in early stages of this disorder.

8 Conclusion

One of the motive of neuroscience is to find the stages from genetic expression to the condition of brain and its related disorders. The paper addressed the different form of disorders of brain and the role of EEG in diagnosing brain diseases. Methods used to detect these diseases are also discussed. Although this study is primarily oriented toward academic audience and researchers, it is also useful for medical practitioners. It is concluded that there are different brain disorders including brain tumors, brain injuries, neurodegenerative, and mental disorder.

References

1. Brookmeyer, R., Johnson, E., Ziegler-Graham, K, & Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement 3*, 186–191. https://doi.org/10.1016/j.jalz.2007.04.381. PubMed: 19595937. [PubMed] [CrossRef].
2. Siuly Siuly, Yanchun Zhang, (2016) Medical Big Data: Neurological Diseases Diagnosis Through Medical Data Analysis. Data Science and Engineering 1 (2):54-64
3. Qiu, Y., et al. (2018). Denoising sparse autoencoder-based ictal EEG classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26*(9), 1717–1726.
4. Gotman, J. (1982). Automatic recognition of epileptic seizures in the EEG. *Electroencephalography and clinical Neurophysiology, 54*(5), 530–40.
5. Kulkarni, N. (2018). Use of complexity based features in diagnosis of mild Alzheimer disease using EEG signals. *International Journal of Information Technology, 10*(1), 59–64.
6. Kaur, B., Dinesh, S., & Roy. P. P. (2017). A novel framework of EEG-based user identification by analyzing music-listening behavior. *Multimedia tools and applications, 76*(24), 25581–25602.
7. Kumar, T. S., Vivek K., & Ram B. P.. (2015). Classification of seizure and seizure-free EEG signals using local binary patterns. *Biomedical Signal Processing and Control, 15*, 33–40.
8. Brunner, C., Birbaumer, N., Blankertz, B., et al. (2015). BNCI horizon 2020: Towards a roadmap for the BCI community. *Brain Computer Interfaces, 2*(1), 1–10.
9. Bao, F. S., Lie, D. Y. -C. & Zhang, Y. (2008). A new approach to automated epileptic diagnosis using EEG and probabilistic neural network. In *2008 20th IEEE International Conference on Tools with Artificial Intelligence.* Vol. 2. IEEE.
10. Narechania, A., Garic, I., Sen-Gupta, I., Macken, M., Gerard, E., & Schuele, S. (2013). Assessment of a quasi-piezoelectric mattress monitor as detection system for generalised convulsions. *Epilepsy And Behaviour, 28,* 172–176.
11. Adrian J. Fowle, Colin D. Binnie, (2000) Uses and Abuses of the EEG in Epilepsy. Epilepsia 41 (s3):S10-S18

12. Guo, Kairui, et al. "EEG-based emotion classification using innovative features and combined SVM and HMM classifier." 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017
13. Wajid Mumtaz, Likun Xia, Syed Saad Azhar Ali, Mohd Azhar Mohd Yasin, Muhammad Hussain, Aamir Saeed Malik, (2017) Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD). Biomedical Signal Processing and Control 31:108-115
14. Jisu Elsa Jacob, Gopakumar Kuttappan Nair, Ajith Cherian, Thomas Iype, (2019) Application of fractal dimension for EEG based diagnosis of encephalopathy. Analog Integrated Circuits and Signal Processing 100 (2):429-436
15. Arnedo, V., Carlson, C., & Devinsky, O. (2009). Detecting nocturnal convulsions: efficacy of the MP5 monitor. *Seizure, 18*(3), 225–227.
16. Gotman, J. (1990). Automatic seizure detection: improvements and evaluation. *Electroencephalography and clinical Neurophysiology, 76*(4), 317–324.
17. Hao, Q., & Gotman, J. (1993). Improvement in seizure detection performance by automatic adaptation to the EEG of each patient. *Electroencephalography and clinical Neurophysiology, 86*(2), 79–87.
18. Cho, D., et al. (2016). EEG-based prediction of epileptic seizures using phase synchronization elicited from noise-assisted multivariate empirical mode decomposition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25*(8), 1309–1318.
19. Maksimenko, V. A., et al. (2018). Artificial neural network classification of motor-related eeg: An increase in classification accuracy by reducing signal complexity. *Complexity 2018.*
20. Jacob, J. E., et al. (2019). Application of fractal dimension for EEG based diagnosis of encephalopathy. *Analog Integrated Circuits and Signal Processing, 100*(2), 429–436.

Red Lesion Detection in Retinal Fundus Image for Diagnosis of Diabetic Retinopathy

M. L. Jadhav, M. Z. Shaikh, and V. M. Sardar

Abstract Diabetes retinopathy is a consequence of diabetes mellitus. It is one of the major reasons of avoidable blindness in rural as well as urban areas. This disease is mainly caused if body is not able to produce enough insulin, causing high blood sugar levels. The minute vessels in the retinal areas get swollen due to high sugar levels in blood, resulting them to rapture and spread of blood in the retinal region. This results in to haemorrhages, i.e. red lesions in retina. With the advancements in the field of image processing, these lesions are easily detected at earlier stage to avoid further consequences. In this research work, the morphological approach has been developed to detect these lesions. This mainly includes pre-processing, detection and elimination of fovea and blood vessels. The microaneurysms being minute lesions are eliminated to detect haemorrhages. The proposed work is carried out on Indian diabetic retinopathy image dataset (IDRiD). The accuracy, sensitivity and specificity achieved with this approach is noted as 98.22%, 75.3% and 91.65%, respectively.

Keywords Diabetic retinopathy · Haemorrhages · Fovea · Blood vessels · Fundus image · Morphology

1 Introduction

As per the World Health Organization (WHO) diabetes has become a rapidly increasing health threat due to sudden rise in blood sugar levels. Long-term diabetes causes many health threats including neuropathy, nephropathy, heart disease and

M. L. Jadhav (✉)
Department of E&TC Engineering, MET's Institute of Engineering, Nasik, M.S, India
e-mail: manisha.shinde1@gmail.com

M. Z. Shaikh
Shree Bhagubhai Mafatlal Polytechnic, Mumbai, M.S, India
e-mail: mzshaikh2003@yahoo.co.in

V. M. Sardar
Department of E&TC Engineering, J.S. College of Engineering, Hadapsar, Pune, M.S, India
e-mail: vijay.sardar11@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_33

retinopathy [1]. It has been estimated in [2] that the occurrence of diabetes was about 2.8% in the year 2000 and may increase to 4.4% till the year 2030 for all age groups. The diabetic retinopathy is result of diabetes, in which the extra pressure on retinal blood vessels is caused due to increased blood sugar levels. This slowly damages the tiny vessels causing them to get swollen and sack-like structure, normally earliest sign of diabetic retinopathy termed as microaneurysms [3]. In addition to this, the blood vessels wall starts weakening, causing haemorrhages [4]. As per the international classification of severity levels of diabetic retinopathy, the microaneurysms are normally visible at early stage of the disease and haemorrhages tend to appear in the following stages. As the disease progresses in later stages, the number of haemorrhages goes on increasing which becomes sight threatening.

The retinal damage can be avoided by regular eye examination followed by proper treatment. The retinal image is captured with the help of fundus camera showing red and bright lesions in retina. The red lesions are microaneurysms and haemorrhages and bright lesions are exudates. Figure 1 shows the retinal image captured with fundus camera with different lesions.

This paper concentrates on detection of red lesions being very important sign to decide the severity level of diabetic retinopathy. The proposed approach attempts contrast limited adaptive histogram equalization (CLAHE) to overcome the problem of non-uniform illumination as well as low contrast and noise in the image. Further morphological operations are applied to detect haemorrhages. The paper is organized as follows: The related literature has been explored in Sect. 2. Section 3 describes the implementation of proposed methodology for pre-processing followed by (CLAHE). Further, the fovea and blood vessels are detected and eliminated to detect haemorrhages. Sections 4 and 5 describe the experimental results and discussion, respectively. The conclusion drawn based on results is presented in Sect. 6.

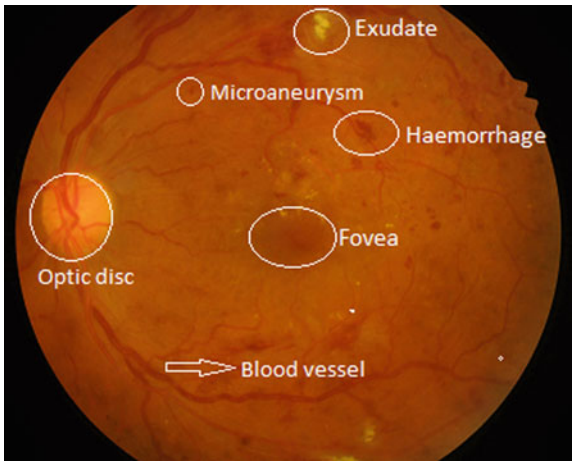


Fig. 1 Retinal image with different types of lesions

2 Related Literature

There are various systems available in the literature for segmentation of DR lesions for an accurate diagnosis.

Different techniques for haemorrhage detection have been compared in [5] and the classifier performance was tested based on sensitivity and specificity. The localization and segmentation of optic disc, vasculature and fovea have presented for different datasets in [6]. A morphological approach for detecting DR lesions is adopted in [7]. The microaneurysms, exudates and haemorrhages were classified using support vector machine (SVM) classifier. The accuracy achieved was 82% and sensitivity and specificity was 86%. The template matching approach with normalized cross correlation along with region growing was used to extract haemorrhage and computing their size respectively. The sensitivity for haemorrhage detection was 85% with four false positives per image [8]. The filtering and SVM approach to detect and classify the red lesions have put forward in [9]. The sensitivity, specificity and accuracy achieved reported were 96.42%, 100% and 96.62%, respectively. The study in [10] has proposed a new method for gamma correction to RGB plane of colour image. Further, the extending histogram of each plane haemorrhages was detected, and the false positives were eliminated changing brightness using nonlinear curve with brightness values of HSV space. The sensitivity and specificity obtained were 80% and 88%, respectively. The random forest classifier to classify DR lesions and nonlesions have put forth in [11]. The candidates were calculated using regional minimas and dynamic features as a function of intensity, after shade correction, contrast enhancement and de-noising. They have achieved a global score over a FROC curve of 0.393. The proposed adaptive threshold technique to detect dark and bright lesions in diabetic retinopathy in [12] found successful. The statistical features were used to segment abnormalities, associated with DR. The study reported performance in terms of sensitivity/specificity for bright and red lesions as 97/89% and 94.2/84.5%, respectively. The classification of normal images from DR images achieves an average sensitivity/specificity of 93.90/76.49%.

3 Methodology

The proposed approach for red lesion detection has been presented in different stages elaborated and depicted in Fig. 2. The IDRiD dataset [13] is used in the research work for implementation consists of 81 images of size of 4288×2848. The ground truths for different lesions of diabetic retinopathy along with grading ground truths are provided along with dataset.

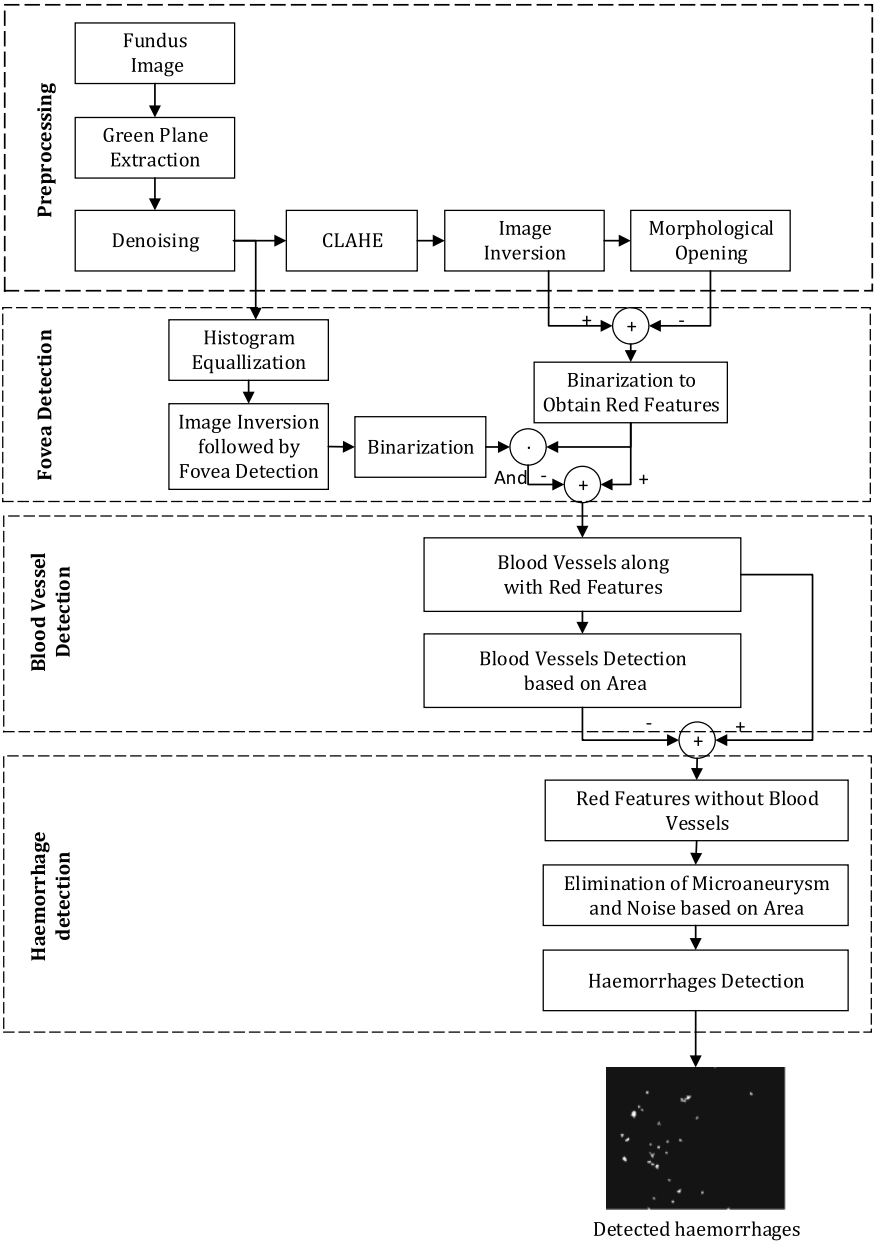


Fig. 2 Schematic representation of the proposed haemorrhage detection system

3.1 Pre-processing

The retinal colour fundus image is composed of red, green and blue plane. The image is resized to 576 × 720 pixels in order to make it suitable for different datasets as well as to reduce the computation time. As the green plane carries good contrast details in the image as compared to red and blue plane, so it is selected for further processing. The median filter of size 3 × 3 is used to remove noise from the green plane image in [14]. The contrast limited adaptive histogram equalization (CLAHE) is applied with clip limit of 0.09 on the filtered image. As CLAHE operates on small regions called as tiles [15], enhancing the contrast of individual section and then tiny sections are integrated using bilinear interpolation to remove the synthetically induced boundaries. Also, it limits the over strengthening of noise by limiting contrast of each uniform section.

The CLAHE operates on filtered image is expressed as

$$f_h = \text{CLAHE}(f_i) \tag{1}$$

where f_i is median filtered image, f_h equalized image is further complemented to make the haemorrhages clearly visible being dark red lesion in colour fundus image.

The complemented image f_{ch} is opened by morphological operation with a disc structuring element with radius of 15 to generate the background image.

$$f_{ch\_\text{open}} = (f_{ch} \ominus B) \oplus B \tag{2}$$

where B is disc-shaped structuring element. f_{ch} is complemented image.

The opened image is further subtracted from complemented image.

$$f_{rh} = f_{ch} - f_{ch\_\text{open}} \tag{3}$$

f_{rh} is image with red candidates detection. Otsu's thresholding is applied on f_{rh} [16], to reduce interclass variance of thresholded black and white pixels, and the image is binarized.

3.2 Fovea Detection

The inverted image after histogram equalization is considered to detect fovea. As fovea is a red colour structure similar in intensity as that of microaneurysms and haemorrhages, so it needs to be detected and omitted. The image is binarized using Otsu's thresholding such that foreground is separated from background. In foreground, other components like blood vessels along with fovea are detected. The

experiment is conducted for different values of connected pixels as 750, 800 and 850 pixels. For 750 connected pixels, some part of fovea is clipped off. Whereas for 850 connected pixels, some part of large blood vessels is added. So the optimum value of connected pixels with area 800 pixels is considered as a part of fovea while removing other components with smaller area. The most circular part among the detected part is considered as fovea for that the compactness measure is applied to detect the fovea [17]. The areas higher than 10,000 connected pixels, being background pixels were removed. Thus, the binary image of fovea is obtained. The logical AND operation is applied on fovea detected image f_{fovea} and red candidate image f_{rh} to generate mask, to remove fovea from the image f_{rh} as given in Eqs. 4 and 5.

Thus, fovea eliminated image f_{f_elim} is obtained by subtracting mask from red candidates image.

$$f_{\text{mask}} = \text{AND}\left(f_{\text{fovea,}}\, f_{rh}\right) \tag{4}$$

$$f_{f\_\text{elim}} = f_{rh} - f_{\text{mask}} \tag{5}$$

where f_{mask} is mask, f_{f_elim} is image with masked fovea.

3.3 Blood Vessel Detection

The resultant image f_{f_elim} contains long blood vessels as well as short blood vessels. The connected components with area more than 1500 pixels are recognized as long vessels; also connected components from the image are extracted based on eccentricity, to find rest of short blood vessels, not connected to long blood vessels due to damages. They are in the eccentricity value from 0.88 to 1. The eccentricity value ranges from zero (circle) and one (line). The image with long blood vessels is denoted as f_{l_b} and the image with short blood vessels is denoted by f_{s_b}. The images with long blood vessels and short blood vessels are ORed as shown in Eq. 6.

$$f_{bv} = \text{OR}(f_{l\_b},\, f_{s\_b}) \tag{6}$$

where f_{bv} is blood vessels image.

3.4 Red Lesion/Haemorrhage Detection

After detection and elimination of fovea and long and short blood vessels, remaining portions are haemorrhages and microaneurysms. The resultant image is obtained subtracting detected blood vessel image from the fovea eliminated image as given in Eq. 7.

$$f_{\text{red}} = f_{f\_\text{elim}} - f_{bv} \tag{7}$$

where f_{bv} is blood vessels image, f_{f_elim} is fovea eliminated image.

Final binary image f_{red} represents red lesions present in the fundus image. In order to remove the microaneurysms and artefacts present in final binary image, connected pixels whose area is less than 30 are removed, resulting into haemorrhages present in the original fundus image. The parameters named true positives, true negatives, false positives and false negatives are used to evaluate the system performance like sensitivity, specificity and accuracy.

4 Experimental Result

The proposed technique of red lesion detection is applied on the IDRiD dataset, the dataset is provided with haemorrhage ground truths. The dataset is available online consisting of 81 images. The proposed technique has correctly detected the haemorrhage regions in retinal image.

It mainly goes through the following steps:

a. Pre-processing
b. Fovea and blood vessels detection and elimination
c. Red lesion/haemorrhage detection.

In pre-processing, to carry out enhancement operation on de-noised green plane CLAHE is used with clip limit 0.09. The morphological opening operation is carried out on complemented image after CLAHE with disc-type structuring element of size 15. The red lesion image is obtained as given in Eq. (3). For detecting fovea region, the connected pixels with area 800 are considered empirically. The detected fovea is removed from red lesion image as given in Eq. (5). The long and short blood vessels are detected and removed using morphological operations. The red lesions image is obtained by suppressing the detected vessels image given by Eq. (7). Finally, the haemorrhages are extracted based on area as shown in Fig. 3g.

The sensitivity and specificity and accuracy are calculated for the detected microaneurysms as follows (Table 1):

$$\text{Sensitivity} = \frac{(\text{True positive})}{(\text{True positive} + \text{False negative})} \tag{8}$$

$$\text{Specificity} = \frac{(\text{True negative})}{(\text{True negative} + \text{False positive})} \tag{9}$$

$$\text{Accuracy} = \frac{(\text{True positive} + \text{True negative})}{(\text{True positive} + \text{True negative} + \text{False positive} + \text{False negative})} \tag{10}$$

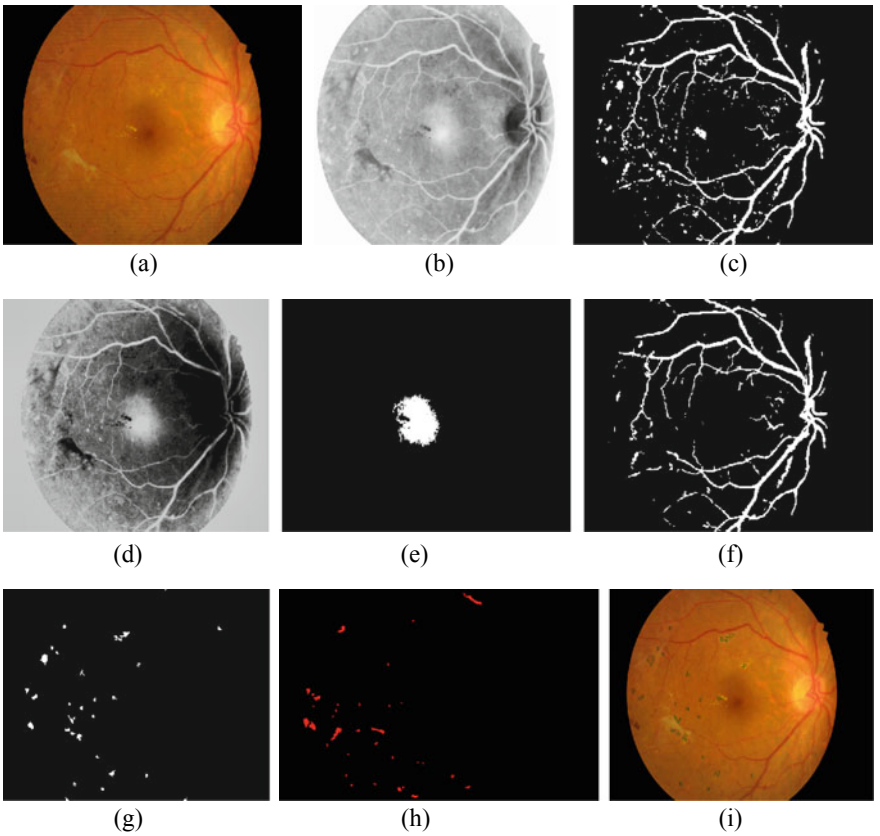


Fig. 3 **a** Original colour fundus image IDRiD-ROC dataset with resolution is of 4288 × 2848 pixels, **b** complemented green plane image after CLAHE, **c** binarized image after Otsu's thresholding, **d** complemented green plane after histogram equalization followed by CLAHE, **e** fovea detected image, **f** detected blood vessels, **g** detected haemorrhages, **h** haemorrhages ground truths, **i** fundus image outlined with detected haemorrhages

Table 1 Average accuracy, sensitivity and specificity calculated for given dataset

No. of images	Accuracy	Sensitivity	Specificity
81 (resolution: 4288 × 2848)	98.22%	75.3%	91.65%

Above results obtained with morphological operations are compared with ground truths provided with source dataset. The accuracy found to be high as 98.22%. Specificity is also appreciable with slight reduction in sensitivity but up to the acceptable mark.

5 Discussion

Red lesion extraction from fundus image is challenging task. Since, contrast may be misinterpreted with background in some cases; the non-uniformity in background intensity may also lead the misinterpretation in the discrimination of lesion and background. In order to overcome this difficulty, simple and effective approach using contrast limited adaptive histogram equalization to enhance contrast of each small region by suppressing noise has been presented in this work. The equalized image shows all bright and red lesions along with retinal features such as optic disc, blood vessels and fovea, whereas the area of interest is only red lesions. So to make bright lesions and optic disc invisible, the image is complimented to make red lesions more prominent.

The fovea detection and elimination are achieved using morphological operations as the contrast of fovea and red lesion is same which may lead to false positives. The long and short blood vessels are detected based on area of connected pixels and by calculating eccentricity, respectively. The addition of detected long and short blood vessels and eliminating it from red features image provides only red lesions present in image. The microaneurysms along with some artefacts were suppressed based on area to detect haemorrhages. With the proposed approach, the accuracy of haemorrhage detection is 98.22% with sensitivity and specificity of 75.3% and 91.65%, respectively.

6 Conclusion

This research work presents the method to detect the red lesions by analysing the fundus image. Being a part of DR detection system, it can be helpful for ophthalmologist to diagnose the disease at earlier stage. For the development of this system, the contrast limited adaptive histogram equalization of image is implemented to overcome the non-uniform illumination in image followed by morphological operations for red lesion detection. The fovea detection is accomplished using optimum value of 800 connected pixels after investigating different settings and maximum value of compactness being nearly circular in shape. Thus, by eliminating fovea, the accuracy is improved by reducing the false positives.

As a future scope, it is proposed to extend the work for reducing the false negatives. False negatives are outcome of misinterpretation due to similarity in intensity of red lesion and background which may be controlled on thresholding while binarization. Moreover, the intensity profile for all images in the dataset may vary. So, the efficient adaptive intensity threshold may give the best performance in terms of sensitivity.

Acknowledgements The authors are thankful to the team members of Eye clinic Nanded (Maharashtra, India) for generating the diabetic retinopathy grading dataset Indian Diabetic Retinopathy Image Dataset (IDRiD) and making it available online.

References

1. World Health Organization. (2019). *Classification of diabetes mellitus.*
2. Wild, S., Roglic, G., Green, A., Sicree, R., & King, H. (2004). global prevalence of diabetes—Estimates for the year 2000 and projections for 2030. *Epidemiology/Health Services/Psychosocial research.*
3. Bharali, P., Medhi, J. P., & Nirmala, S. R. (2015). Detection of hemorrhages in diabetic retinopathy analysis using color fundus images. In 2015 *IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS).*
4. Júnior, S. B., & Welfer, D. (2013). Automatic detection of microaneurysms and hemorrhages in color eye fundus images. *International Journal of Computer Science & Information Technology (IJCSIT), 5*(5). https://doi.org/10.5121/ijcsit.2013.5502.
5. Gulati, S., Kleawsirikul, N., & Uyyanonvara, B. (2012). A review on hemorrhage detection methods for diabetic retinopathy using fundus images. *International Journal of Biological, Ecological and Environmental Sciences (IJBEES), 1*(6). ISSN 2277-4394.
6. Prentasic, P. (2013). *Detection of diabetic retinopathy in fundus photographs,* (pp. 6–8). University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia.
7. Acharya, U. R., Lim, C. M., Ng, E. Y. K., Chee, C., & Tamura, T. (2009). Computer based detection of diabetes retinopathy stages using digital fundus images. *Journal of Engineering in Medicine, 223*(H5), 545–553.
8. Bae, J. P., Kim, K. G., Kang, H. C., Jeong, C. B., Park, K. H., & Hwang, J. -M. (2011). A study on hemorrhage detection using hybrid method in fundus images. *Journal of Digital Imaging, 24*(3), 394–404.
9. Mane, V. M., Kawadiwale, R. B., & Jadhav, D. V. (2015). Detection of red lesions in diabetic retinopathy affected fundus images. In *2015 IEEE International Advance Computing Conference (IACC)*, Banglore (pp. 56–60).
10. Hatanaka, Y., Nakagawa, T., Hayashi, Y., Kakogawa, M., Sawada, A., Kawase, K., et al. (2008). Improvement of automatic hemorrhages detection methods using brightness correction on fundus images. *Proceedings of SPIE, 6915,* 69153E.
11. Séoud, L., Faucon, T., Hurtut, T., Chelbi, J., Cheriet, F., & Langlois, J. M. P. (2014). Automatic detection of microaneurysms and haemorrhages in fundus images using dynamic shape features. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI 2014).*
12. Sengar, N., & Dutta, M.K. (2017). Automated method for hierarchal detection and grading of diabetic retinopathy. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization.* https://doi.org/10.1080/21681163.2017.1335236.
13. Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., & Meriaudea, F. (2018). Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research. Dataset: https://doi.org/10.21227/h25w98, Dataset License: CC-BY 4.0, Data 2018, 3, 25; https://doi.org/10.3390/data3030025, www.mdpi/journal/data.
14. Vikhe, P. S., & Thool, V. R. (2019). A wavelet and adaptive threshold-based contrast enhancement of masses in mammograms for visual screening. *International Journal of Biomedical Engineering and Technology, 30*(1), 2019.
15. Dash, J., & Bhoi, N. (2018). Retinal blood vessel segmentation using Otsu thresholding with principal component analysis. In *Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018).* IEEE Xplore Compliant—Part Number: CFP18J06-ART, ISBN 978-1-5386-0807-4; DVD Part Number: CFP18J06DVD, ISBN 978-1-5386-0806-7.
16. Yousef, J. (2015). Image binarization using Otsu thresholding algorithm. https://doi.org/10.13140/rg.2.1.4758.9284.
17. Wirth, M. A. (2004). *Shape analysis & measurement*. Ph.D. University of Guelph Computing and Information Science image Processing Group.

Websites

18. https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid.

Automated Detection and Counting of Red-Dots from Digital Fundus Images

Parashuram Bannigidad and Asmita Deshpande

Abstract Diabetic Retinopathy is an eye complication that occurs among patients suffering from type 1 and type 2 diabetes. The earliest clinical features of Non-proliferative Diabetic Retinopathy are presence of very small Red-dots having size less than 125 Microns. These Red-dots may be Microaneurysms and/or Haemorrhages. In this paper, morphological operations and thresholding are prominently used for detection of Red-dots. LBP features are extracted and decision tree classifier is used to distinguish between healthy and diseased images. The Red-dots may appear and disappear on retinal surface of a diabetic patient during the course of treatment. Red-dot counting has been suggested as an appropriate marker of understanding the progression of Diabetic Retinopathy. Hence, counting of Red-dots is also an important module of the proposed technique. The proposed method was implemented on 608 images from e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases. The accuracy values for e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases are 0.93, 1.0, 0.97, 0.98 and 1.0 respectively. The encouraging results obtained in this experiment testify the efficiency of the proposed method.

Keywords Red-dots · Microaneurysms · Haemorrhages · Digital fundus images · Thresholding · LBP features · Decision tree classifier

1 Introduction

Diabetic Retinopathy is an eye complication that occurs among patients suffering from type 1 and type 2 diabetes. In the early stages, a patient suffering from Diabetic Retinopathy may not suffer any problems related to vision. However, as the disease progresses, it can cause severe damage to the eye and eventually lead to blindness.

P. Bannigidad (✉) · A. Deshpande
Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India
e-mail: parashurambannigidad@gmail.com

A. Deshpande
e-mail: asd_bca@yahoo.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_34

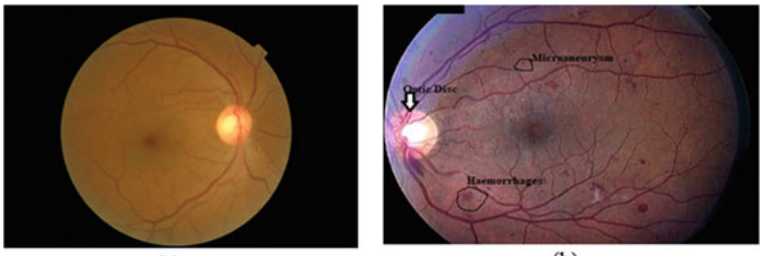


Fig. 1 **a** Normal fundus image, **b** Fundus image showing moderate NPDR

Diabetic Retinopathy is generally divided into four stages: Mild Non-proliferative Diabetic Retinopathy, Moderate Non-proliferative Diabetic Retinopathy, Severe Non-proliferative Diabetic Retinopathy and Proliferative Diabetic Retinopathy. The earliest clinical features of Non-proliferative Diabetic Retinopathy (NPDR) are presence of very small Red-dots having size less than 125 Microns. These Red-dots may be Microaneurysms and/or Haemorrhages. Ophthalmologists depend on digital fundus photographs to diagnose retinal diseases. A normal fundus image is shown in Fig. 1a and Fig. 1b shows fundus image having moderate Non-proliferative Diabetic Retinopathy.

A number of researchers have developed algorithms for automated detection of Diabetic Retinopathy. Arnay et al. [1] developed a technique based on an Ant Colony Optimization for segmentation of optic cup in which a heuristic is generated based on the optic disc intensity gradient and curvature of blood vessels. Kaur and Mittal [2] worked on segmentation of blood vessels in fundus images based on vascular and non-vascular regions and joined the candidate vessel structures to create connectivity. Habib et al. [3] have proposed a method to find candidate Microaneurysms with Gaussian matched filter and then remove the false positives from the dataset using classifier. Cao et al. [4] extracted 25 $\times$ 25 pixel patches from fundus images, implemented principle component analysis and tested various classifiers. To deal with false detections on blood vessels, novel filters have been proposed which can distinguish between red lesions and blood vessels. Shrivastav et al. [5] proposed a grid-based method for detection of lesions using multiple kernel learning. Shan and Li [6] presented a stacked sparse autoencoder which works on small image patches and their intensities to detect Microaneurysms. Zhou et al. [7] developed Haemorrhage detection method by adopting dual-scale segmentation for dark regions. Jose et al. [8] explored a template matching method and morphological operations for detection of important anatomical structures of a retina. Amin et al. [9] reviewed various methods for detection of lesions, exudates and other symptoms that characterize Non-proliferative Diabetic Retinopathy. Parashuram Bannigidad and Asmita Deshpande [10, 24, 25] explored a multi-stage approach for detection of exudates based on thresholding and morphological operations. The main objective of this paper is to detect the Red-dots comprising Microaneurysms and Haemorrhages from various public fundus databases as well as develop a method for counting them.

Table 1 Data description

S. No.	Database name	Acquisition	FOV	Image size	Format	H	D	T
1	e-Ophtha	OPHDIAT	50°	1696 × 2544	jpg	233	148	381
2	HRF	Canon CR-1	45°	2336 × 3504	jpg	00	18	18
3	DIARETDB1	Fundus camera	50°	1152 × 1500	png	05	84	89
4	MESSIDOR	TopCon TRC NW6	45°	1488 × 2240	tiff	32	68	100
5	STARE	TopCon TRV 50	45°	700 × 605	jpg	10	10	20

H-Healthy images, D-Diseased images, T-Total images

Counting the Red-dots allows ophthalmologists to track the progress of the disease [11] and provide necessary treatment.

2 Materials and Methods

To execute the proposed method, digital fundus images from five different data sets, namely e-Ophtha [12], High Resolution Fundus (HRF) [13], DIARETDB1 [14], MESSIDOR [15] and STARE [16] have been used. Table 1 describes the fundus images that makeup the data sets.

3 Proposed Methodology

The proposed method is based on implementing the digital image processing techniques such as image enhancement, segmentation, feature extraction and classification.

3.1 Preprocessing

The aim of preprocessing is to resize the image to some standard size, adjust the contrast and highlight the pathological symptoms of Diabetic Retinopathy like Microaneurysms and Haemorrhages. Figure 2a shows the RGB fundus image, Fig. 2b shows image after green channel extraction. Similarly, Fig. 2c shows the grey scale

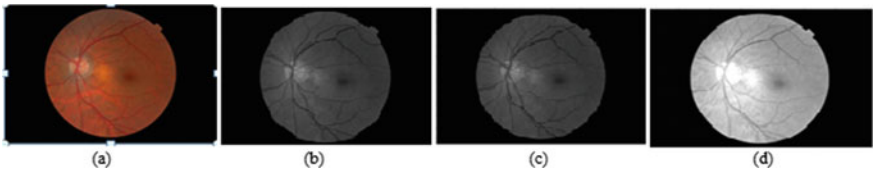


Fig. 2 **a** Fundus image, **b** green channel, **c** grey scale, **d** contrast adjusted image

image and Fig. 2d shows the image after contrast enhancement. The algorithmic steps for preprocessing are given below:

Algorithm

Step 1: Read RGB Digital fundus image I
Step 2: Resize I to 576 X 720 pixels
Step 3: Extract the green channel from I
Step 4: Convert the image I to grey scale
Step 5: Apply CLAHE to I in order to enhance the contrast.

3.2 Segmentation, Feature Extraction and Classification

Human retina consists of fine blood vessels that supply blood to the eye. It also comprises the optic disc that connects optic nerves to the retina. The blood vessels and optic disc are barriers in accurate detection of Red-dots. Hence, they have to be eliminated. Canny edge detection followed by morphological operations [17] aid in removing the blood vessels. However, the optic disc can be eliminated by computing a circular mask and subtracting it from segmented image. Microaneurysms and Haemorrhages are very small Red-dots present on the surface of the retinal image. Since digital fundus images exhibit dissimilarities in intensity and texture; segmentation based on thesholding is proposed in this paper. The algorithmic steps for segmentation are described below:

Algorithm

Step 1: Execute canny edge detection to image I obtained after preprocessing
Step 2: Eliminate blood vessels by applying morphological operations to the image I obtained in Step 1
Step 3: Apply column wise neighbourhood operations and thresholding ($T = 0.7$) to the image obtained in Step 2
Step 4: Generate a circular mask M using mesh grid vectors$x$,y and medians of a row and column coordinates
Step 5: Subtract mask M from image obtained in Step 3 to eliminate the optic disc
Step 6: Apply morphological closing operation to the image obtained in Step 5 to eliminate the outer border

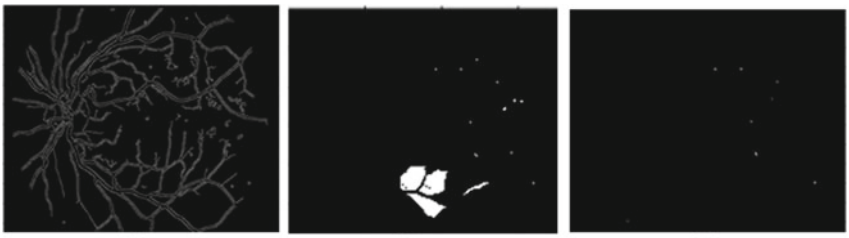


Fig. 3 **a** Image after canny edge detection, **b** optic disc and blood vessels eliminated, **c** image after segmentation of Red-dots

Step 7: Apply contrast limited adaptive histogram equalization to the image obtained in Step 6
Step 8: Complement the image obtained in Step 7
Step 9: Invert the image obtained in Step 7
Step 10: Apply AND operation to image obtained in Step 8 and 9 to extract the Red-dots accurately.

Figure 3a shows the fundus image after canny edge detection. Figure 3b depicts the image after optic disc and blood vessel elimination. Figure 3c shows the segmented image after Red-dots are detected. Once the fundus images have been segmented accurately, local binary pattern (LBP) features are extracted. These features encode local texture information that can facilitate tasks such as classification of symptoms that characterize Non-proliferative Diabetic Retinopathy. Further, the decision tree classifier is applied to distinguish between healthy and diseased images.

4 Red–Dot Counting

The clinical symptoms of Diabetic Retinopathy vary in every patient. The Red-dots may appear and disappear during the course of treatment. Disappearance generally occurs due to closure of capillaries. Red-dot counting has been suggested as an appropriate marker for understanding the progression of Diabetic Retinopathy [18]. This can aid ophthalmologists during follow-up treatment of a retinopathy patient. It can also be used to establish the retinopathy grade as, "NORMAL", "MILD", "MODERATE" and "SEVERE" depending on the number of Red-dots [18]. In the proposed method, Red-dots counting is initiated after segmentation process. This is realized by finding connected components with 8–connectivity and counting them. These connected components indicate Microaneurysms and Haemorrhages. The following algorithmic steps are used for counting the Red-dots:

Algorithm

Step 1: Read segmented Image I

Table 2 Red-dots counting and grading of images from e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE

Database	Normal	Mild	Moderate	Severe
e-Ophtha	66	106	121	88
HRF	00	02	01	15
DIARETDB1	01	18	27	43
MESSIDOR	05	47	18	30
STARE	00	02	07	11

Step 2: Initialize count = 0
Step 3: Find connected components with 8–connectivity
Step 4: Increment count by 1
Step 5: Repeat steps 3 and 4 for all connected components
Step 6: if count = 0 then GRADE = "NORMAL"
Step 7: if count > 0 and count <=5 then GRADE = "MILD"
Step 8: if count > 5 and count <=15 then GRADE = "MODERATE"
Step 9: if count > = 15 then GRADE = "SEVERE"

The results of Red-dots counting after segmentation are shown in Table 2.

5 Experimental Results and Discussions

The proposed method was implemented on publicly available fundus data sets e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE. The results of classification of digital fundus images from these databases using LBP features with decision tree classifier are shown in Table 3.

Accuracy, recall and precision are three important measures of performance used to evaluate the proposed algorithm. In image processing and pattern recognition, accuracy represents how many images the classifier could identify correctly as either

Table 3 Classification of fundus images from e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases using LBP features and decision tree classifier

Database	Classification results			
	TP	FP	FN	TN
e-Ophtha	128	06	20	227
HRF	15	00	00	03
DIARETDB1	84	02	0	03
MESSIDOR	67	01	01	31
STARE	10	00	00	10

TP-True Positives, FP-False Positives, FN- False Negatives, TN- True Negatives

Table 4 The values of performance evaluation measures on e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases using LBP features

Database	Decision tree classifier		
	Accuracy	Recall	Precision
e-Ophtha	0.93	0.86	0.95
HRF	1.00	1.00	1.00
DIARETDB1	0.97	1.0	0.97
MESSIDOR	0.98	0.98	0.98
STARE	1.00	1.00	1.00

Table 5 Comparative results of proposed method with other methods

Database	Author	Accuracy
e-Ophtha	Ren et al. [19]	AUC 0.84
	Proposed Method	0.93
HRF	Bharali et al. [20]	0.98
	Proposed Method	1.0
DIARETDB1	Sreng et al. [21]	0.90
	Proposed Method	0.97
MESSIDOR	Bharali et al. [22]	0.98
	Proposed Method	0.98
STARE	Arati et al. [23]	0.84
	Proposed Method	1.00

healthy or diseased from a given dataset, whereas the ability of a classifier to find all positive images correctly is known as recall. However, in information retrieval systems, precision is the fraction of relevant instances of images classified among the retrieved instances. The values of performance evaluation measures on e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases are summarized in Table 4.

When the proposed method was tested with the Decision tree classifier on e-Ophtha database, it yielded values of accuracy, recall and precision 0.93, 0.86, 0.95. For HRF database, it yielded accuracy 1.0, recall 1.0 and precision 1.0. Similarly, the values of accuracy, recall and precision for DIARETDB1, MESSIDOR and STARE databases are 0.97, 1.0, 0.97, 0.98, 0.98, 0.98 and 1.0, 1.0, 1.0 respectively. The comparative results of the proposed algorithm using LBP features with other methods suggested by various researchers are given in Table 5.

6 Conclusion

The earliest signs of Non-proliferative Diabetic Retinopathy are found on retinal surface as tiny Red-dots. This paper proposes an automated method for segmenting the Red-dots using morphological operations and thresholding. Ophthalmologists

can track the appearance and disappearance of Red-dots to monitor the disease. Hence, this paper also explores counting the segmented structures to check the progress of the disease. The proposed method extracts LBP features and applies decision tree classifier to distinguish between healthy and diseased images. The proposed technique was implemented on 608 images from e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases. The accuracy values for e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases are 0.93, 1.0, 0.97, 0.98 and 1.0 respectively. The encouraging results obtained in this experiment testify the efficiency of the algorithm.

Acknowledgements The authors wish to acknowledge the developers of e-Ophtha, HRF, DIARETDB1, MESSIDOR and STARE databases for facilitating the research fraternity with fundus datasets publicly and freely.

References

1. Arnay, R., Fumero, F., & Sigut, J. (2017). Ant colony optimization-based method for optic cup segmentation in retinal images.*Applied Soft Computing,* 409–417.
2. Kaur, Jakirat, & Mittal, Deepti. (2017). A generalized method for the detection of vascular structure in pathological retinal images. *Biocybernetics and Biomedical Engineering, 37,* 184–200.
3. Habib, M. M., Welikala, R. A., Hoppe, A., Owen, C. G., Rudnicka, A. R., & Barman, S. A. (2017). Detection of microaneurysms in retinal images using an ensemble classifier. *Informatics in Medicine Unlocked, 9,* 44–57.
4. Cao, W., Czarnek, N., Shan, J., & Li, L. (2017). Microaneurysm detection in fundus images using small image patches and machine learning methods. In*IEEE International Conference on Bioinformatics and Biomedicine,* USA (pp. 1–8).
5. Srivastavaa, R., Duana, L., Wonga, D. W. K., Liua, J., & Wong, T. Y. (2017). Detecting retinal microaneurysms and hemorrhages with robustness to the presence of blood vessels.*Computer Methods and Programs in Biomedicine,138*, 83–91.
6. Shan, J., Li, L. (2016). A deep learning method for microaneurysm detection in fundus images. In*IEEE First Conference on Connected Health: Applications, Systems and Engineering Technologies* (pp. 357–358).
7. Zhou, L., Qiao, Y., & Yang, J. (2016). Automatic hemorrhage detection in color fundus images based on gradual removal of vascular branches. In *IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ (pp. 399–403).
8. Molina-Casado, J. M., Carmona, E. J., & García-Feijoó, J. (2017). Fast detection of main anatomical structures in digital retinal images based on Intra and inter structural relational knowledge.*Computer Methods and Programs in Biomedicine,149*, 55–68.
9. Amin, J., Sharif, M., & Yasmin, M. (2016). A review on recent developments for detection of diabetic retinopathy.*Scientifica (Cairo)*, 1–20.
10. Bannigidad, P., & Deshpande, A. (2019). Exudates detection from digital fundus images using GLCM features and decision tree classifier. In K. C. Santosh and R. S. Hegadi (Eds.), *Recent trends in image processing* (pp. 245–257), Springer.
11. Cunha-Vaza, J., Bernardes, R. (2005). Non proliferative retinopathy in diabetes type 2. Initial stages and characterization of phenotypes.*Progress in Retinal and Eye Research,24*, 355–377.
12. Decencière, G. C., Zhang, X., Thibault, G., Klein, J. -C., Meyer, F., Marcotegui, B., et al. (2013). TeleOphta: Machine learning and image processing methods for teleophthalmology. In*IRBM* (pp. 196–203).

13. Köhler, T., Budai, A., Kraus, M., Odstrcilik, J., Michelson, G., & Hornegger, J. (2013). Automatic no-reference quality assessment for retinal fundus images using vessel segmentation. In*Proceedings of 26th IEEE International Symposium on Computer-Based Medical Systems* (pp. 95–103).
14. DIARETDB1 Home page: http://www.it.lut.fi/project/imageret/diaretdb1/.
15. Decencière et al. Feedback on a publicly distributed database: the Messidor database.*Image Analysis & Stereology,33*(3), 231–234.
16. STARE Homepage: http://cecas.clemson.edu/~ahoover/stare/.
17. Gonzalez, R. C., Woods, R. E. (2014).*Digital image processing* (3rd ed.).
18. Torrent-Solans, T., Duarte, L., Aleida, E., Bernardes, R., Cunha-Vaz, J. Red dots counting on digitalized fundus images of mild non-proliferative retinopathy in diabetes type 2.*Investigative Ophthalmology & Visual Science*, *2985*.
19. Ren, F., Cao, P., Li, W., Zhao, D., & Zaiane, O. (2017). Ensemble based adaptive oversampling method for imbalanced data learning in computer aided detection of Microaneurysm. *Computerized Medical Imaging Graphics, 55,* 54–67.
20. Bharali, P., Medhi, J. P., & Nirmala, S. R. (2015). Detection of haemorrhages in diabetic retinopathy analysis using color fundus images. In*Proceedings of IEEE 2nd International Conference on Recent Trends in Information Systems* (pp. 237–242).
21. Sreng, S., Maneerat, N., & Hamamoto, K. (2017). Automated microaneurysms detection in fundus images using image segmentation. In*Proceedings of the International Conference on Digital Arts, Media and Technology (ICDAMT)*, Chiang Mai, Thailand, 1–4 March 2017 (pp. 19–23).
22. Khojasteh, P., Aliahmad, B., & Kumar, D. K. (2018). Fundus images analysis using deep features for detection of exudates, haemorrhages and Microaneurysms. *BMC Opthalmology, 18,* 1–13.
23. Manjaramkar, A., & Kokare, M. (2018). Statistical geometrical features for microaneurysm detection.*Journal of Digital Imaging*, 224–234.
24. Bannigidad, P., & Deshpande, A. (2018). A multistage approach for exudates detection in fundus images using texture features with k-NN classifier. *International Journal of Advanced Research in Computer Science, 9*(1), 1–5.
25. Bannigidad, P., & Deshpande, A. (2018). Exudates detection in digital fundus images using GLCM features with SVM classifier. *International Journal of Modern Electronic and Communication Engineering, 6*(6), 184–189.

IoT-Based Patient Health Monitoring: A Comprehensive Survey

B. Sridhara Murthy and Suresh K. Peddoju

Abstract Nowadays, health monitoring is one of the rising fields that is rapidly creating and gaining quick progress in its advancements. Various analysts have proposed different structures for patient health monitoring as per the technical advancements in healthcare systems. During these days, use of Internet is unlimited; Internet of things (IoT) is among of the risen field as of late in which many have had the choice to merge it into different applications. Normally, IoT is utilized to interconnect distinctive accessible resources of medical and related devices. It offers amazing, successful and productive medicinal services to the patients. Health monitoring is the key for dynamic and assisted living. The ongoing advances in IoT improve the way of life of patient by strongly helping and changing the patient by monitoring his health. In this paper, we provided an overview on best in class research and improvements in patient health monitoring systems.

Keywords Patient health monitoring · IoT · Comprehensive survey · Artificial intelligence · Deep learning

1 Introduction

During last few decades, life span of human across the world has improved due to advancements in healthcare systems and diagnostic technologies. Apart from these developments in healthcare systems, people getting awareness about nutrition food, keeping clean environment and primary care about health [1–4]. Indeed with developments in healthcare monitoring, life expectancy will increase and with that fact aged population will be more in compared with birth rates. As per World Health Organization (WHO), by 2050, the old age population is higher than the age of 14

B. S. Murthy (✉)
Kakatiya Institute of Technology and Science, Warangal, India
e-mail: bsm.cse@kitsw.ac.in

S. K. Peddoju
Florida International University, Miami, USA

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_35

[5]. In addition to this, around 15% of world population will suffer with different disabilities and diseases [6]. In general, people with disabilities will have more health issues due to their limited mobility. The general health issues such as heart diseases, cancer, pneumonia, diabetes and many more are the common chronic diseases which affect many people. Approximately, 70% of deaths in adults are mainly caused due to chronic diseases. Diabetes is another major chronic disease which impact many people. Due to daily activities and food habits, people used to get diabetes and may lead to long-term complications and issues with kidney, blindness and other problems. Apart from this, the population of elderly people is increasing almost in all countries. These people require regular assistance and care for their regular activities and proper healthcare. Hence, lot many elderly caretaking centers and services are evolved and grabbing money from the people. Though, their life span is not increasing due to insufficient care and timely addressing the healthcare problems and issues. The life expectancy will decrease if we do not address chronic diseases and other health-related problems. There is an acute need for healthcare services to increase the life span of human.

Nowadays, tremendous developments are taking place in healthcare services. Remarkable advancements in technologies and tools are evolving for patient healthcare over last five years. Though the cost of these present-day healthcare services keep on rise due to developments and usage of new diagnostic tools as well as clinical treatment. A significant spending on healthcare services will change socio-economic structure of almost all countries in the world. Most of the countries proposed good amount of budget for healthcare services and wellbeing of people. In view of these, there should be better healthcare and monitoring services in low cost, efficient and affordable range.

So as to adapt to the developing requirement for patients medicinal services and administrations, it is basic to create moderate, unpretentious and simple to utilize healthcare solutions. Nowadays at home, people have been using natural and wearable devices [7] to monitor real-time information of patients remotely with minimal cost. Smart homes may enable the patients to remain in their agreeable home conditions rather than costly and restricted healthcare facilities. Healthcare personnel can likewise monitor the general health state of the patients continuously and give their opinion and further actions need to be done.

In this paper, we present a state-of-the-art research and developments in patient healthcare monitoring and how current technologies are addressing issues related to patient healthcare by monitoring vital signs such as blood pressure, heart rate, temperatures, pulse rate and other body parameters. The main objective of this paper is to discuss about different contributions of researchers in the development of patient healthcare monitoring systems.

Further, this paper is organized into three sections. Section 2 discusses about general architecture of IoT-based patient health monitoring system and their characteristics. The state-of-the-art research is discussed in related work of Sect. 3. Conclusion of paper is presented in Sect. 4.

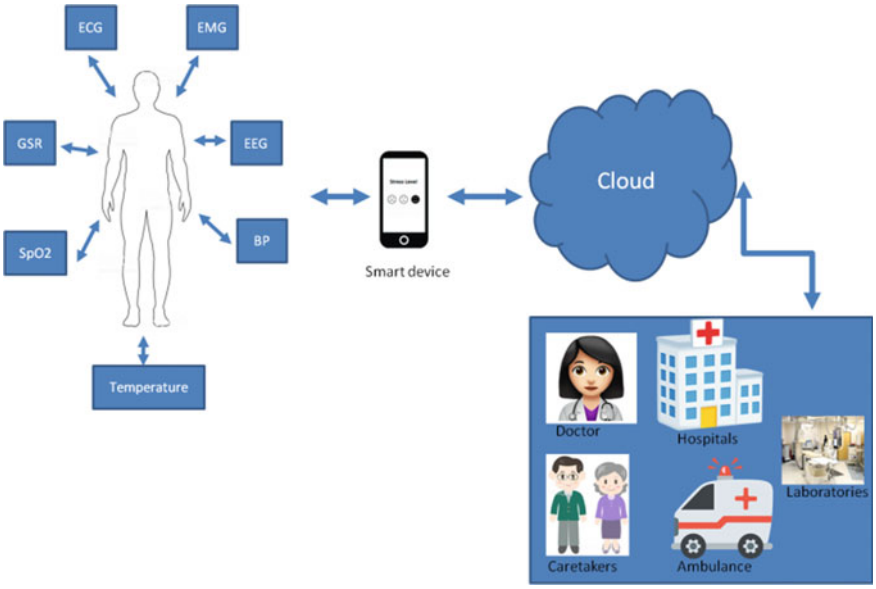


Fig. 1 General architecture of patient healthcare monitoring system

2 Patient Health Monitoring

Nowadays, patient's health is taken care by different healthcare monitoring systems. These systems adopt recent technologies such as IoT, cloud and wearable devices to monitor the status of patient in real time [8]. From the decade, vast developments are taken place in wearable technologies to extract physiological parameters such as ECG, EEG, EFG, BP, temperature, SpO2, GSR and many more. The general architecture of patient monitoring system is depicted in Fig. 1. It shows how patient physiological parameters are transmitted to cloud through smart devices, and there after, they will be processed, assessed and intimated to the respective people and departments linked to the concerned patient.

3 Related Work

The quality of life is major concern of the people and they are looking for new technologies and innovative healthcare solutions to address the needs of people. In current day scenario, Internet of things (IoT) is playing a vital role in providing good healthcare solutions [9]. IoT is connecting all the objects in our daily life including patient health-related objects such as physiological monitoring devices. To keep people active and healthy, a promptly available present-day healthcare system is

demonstrating to be viable in sparing costs, diminishing sickness and drawing out life. There is a need for enhanced healthcare monitoring system to offer better services to the patients with IoT-based technology and easily assess their health status through smart phones, wearables and other devices.

Recent years have seen an ascending in wearable sensors, and today, a few gadgets are monetarily accessible for activity-based patient healthcare. An ongoing healthcare framework should give better healthcare services to individuals whenever at anyplace in a moderate and patient friendly way. As of now, the healthcare services framework going to transform from a conventional way to patients focused methodology. In the current world scenario, wide usage of portable technologies and shrewd devices plays a major role in monitoring health status. Healthcare providers are exploiting this situation with these advancements though a lot of improvements taking place in wearable and clinical environments to provide better monitoring of patient. Apart from these technologies, various other healthcare monitoring devices are being used in healthcare market such as mobile applications, patient forums and ICT-based devices to assist and provide alerts to patients.

Ali et al. [4] proposed an automatic health monitoring system; the functionality of this system is as follows. This system will help residents of all age groups by monitoring automatic voice disorder. In this method, the authors proposed a linear prediction analysis which detects the disorders in voice by determining the source signal from the speech. They compared both normal and disordered subjects. This is a smart healthcare system to help patient suffering from voice complications in smart cities.

Deepasri et al. [10] proposed IoT-based healthcare system to enable the patient to pass on different message to specialists, medical caretaker or his/her friends and family sitting at home or office over the web. The framework utilizes microcontroller base hardware to accomplish this usefulness. It utilizes hand movement acknowledgment circuit and a recipient in addition to transmitter circuit. The given framework is intended to get and process these directions and show them over the LCD display just as transmit the information online over to IoT gecko server. The IoT gecko server at that point shows this data on the web, to accomplish the ideal yield. IoT is used extensively in monitoring patient health condition mainly in case of chronic diseases such as diabetes, pneumonia, asthma and many more. Gómez et al. [11] proposed a design based on metaphysics to monitor health of patients suffering with incessant infections.

These days healthcare environment has created science and information dependent on wireless-sensing hub technology arranged. Patients are confronting a hazardous circumstance of unexpected death because of the particularly of heart issues and attack which is a result of nonexistence of good restorative support to patients at the required time. This is for extraordinarily checking the old age patients and educating specialists and friends and family. So Krishnan et al. [12] proposing an imaginative venture to evade such unexpected demise rates by utilizing patient health monitoring that utilizes sensor technology and web to convey to the friends and family if there

should arise an occurrence of issues. This framework utilizes temperature and heartbeat sensor for monitoring patient's health. If there should arise an occurrence of any sudden changes in quiet pulse or internal heat level, alarm is sent about the patient utilizing IoT.

Saha et al. [13] considered and defined health as a degree of helpful and metabolic power of a living being. In people, it is the intensity of individuals or communities to adjust and self-oversee once confronting physical, mental, psychological and social changes. Monitoring the health remaining of the patient gathering might be an inconvenient undertaking. The goal of this paper is to monitor patient's health with the help of sensors and internet. The health monitoring framework can monitor patient's heartbeat rate, eco rate of heart, pressure level rate, temperature and so forth.

Abdelgawad et al. [14] presented an IoT framework customized for human services applications. According to the proposed model, the gathered information is transferred to the cloud where it is taken over and separated based on the need of information. The information which is not in use can be sent back by the criticism activities to the client. A model of the proposed architecture has been built and results in this paper show its performance advantages.

Patel et al. [15] proposed IoT-based secure healthcare system for smart hospitals. In this paper, they discussed about healthcare framework and importance of data with security concerns. Distributed computing is the most significant worldview in IT-health. All the medical data of the patient just as the specialist and patient personal data store in neighborhood mode just as cloud, so at whatever point it required the information will be effectively accessible. Understanding medicinal information is put away in framework just as cloud, so malicious attack and undesirable access may cause a destructive to patient health. Security is generally significant and critical piece of healthcare.

The popularity of utilization of mobile phones are increasing day by day particularly in developing nations, they can be utilized for productive medicinal services the executives. In this paper, authors of Kodali et al. [16] proposed a novel framework for improving healthcare system with the help of mobile phones with NFC and Bluetooth interfaces. Apart from this, they also used smart card innovation on tamper resistant secure element (SE) for putting away credentials and secure information.

Dar et al. [17] proposed a smart phone-based accident detection system. One way to deal with killing the postponement between accident event and specialist on call dispatch is to use in-vehicle automatic accident recognition and warning frameworks, which sense when car crashes happen and promptly advise emergency workforce. In this paper, authors target to find the event of any accident and reporting the area of accident to the closest emergency vehicle, with the goal that quick help can be given by rescue vehicle. This framework also gives good result in avoiding and preventing false accident detection.

Wu et al. [18] in their paper investigated the utilization of wearable sensing, smart phone along with video streaming and proposed STREMS which is an effective keen continuous pre-emergency clinic correspondence framework for Emergency Medical Services (EMS). Initially, they presented a wearable physiological sensing solution to help multi-dimensional telemetry monitoring for an emergency vehicle working

at as basic life support, a kind of EMS administration level without complex medical equipment or paramedics. Then they proposed to fabricate cloud-based real-time information sharing platform, empowering automated streaming.

Rahman et al. [19] in their paper proposed an Exoskeleton Arm. It is a battery controlled upper-body automated arm which expands human efficiency and quality. Expanding arm quality by forty pounds, Exoskeleton Arm rehabilitates individuals with back injuries, enabling them to reconstruct muscle and relearn engine control. The exoskeleton innovation additionally helps those lifting objects as a component of their day-by-day work, especially in development or conveyance driven positions.

Uma et al. [20] proposed a model to take care of the issues existing in getting and identifying diabetes decease. This framework applied different data mining techniques such as clustering and classification to diagnose what type of diabetes patient have by transforming patient body details to smart system. They considered and collected 650 patients' data to analyze and identify severity of diabetes such as mild, moderate and severe.

Huge data generated in healthcare industry in different forms with floating of images, videos, text, prescriptions and many more which accumulate as big data. Rishika et al. [21] presented a state-of-the-art to better describe about storage and retrieval methods, different tools and techniques used in big data analytics for healthcare industry and utilization of clouds to better provide the services to the people and patients.

Suresh et al. [22] discussed about new technologies such as big data and cloud to solve the problems arising in healthcare. With the growing data needs, an efficient and timely attempt to give proper solutions will reduce the death rates. Authors try to give a solution to identify diabetes patient's severity with the help of machine learning algorithms. A detailed study on diabetes dataset was made with different machine learning algorithms.

Ajay et al. [23] gave overview of wireless infrastructure and corresponding devices which collect data and transform to cloud and other gateways. This paper especially discusses about state-of-the-art solutions to collect the healthcare data, process it and transform to cloud for further analysis and action. It also discussed about different deployment strategies for healthcare solutions. This chapter describes the state of the art and use cases associated to healthcare monitoring which involves the usage of wireless infrastructure.

Suresh et al. [24] proposed an approach to quickly make precise decisions to doctors by analyzing big data collected from different sources such as blood tests, x-ray, etc. According to the authors, this model also shows information such as previous patient's details, different specialist's opinions and their advices during the same situations with the help of cloud and provides analytics over it. Suresh et al. [25] proposed a framework which aims to provide prompt treatments to the patients using wearable devices and IoT technologies. They proposed a model which also monitors patient health information by using wearable devices and that will disseminate to IoT cloud. Further, health status of the patient is monitored with the help of predictive analytics. The same authors proposed different solutions [26–29] to address patient-related issues and its security concerns.

After extensive study of related literature on patient health monitoring system, till now no work has been described on usage of IoT-based health monitoring system for a patient to monitor overall health statistics and provide timely alerts to patients by predicting patient's health data in real time through deep learning and advanced analytics techniques.

4 Conclusion

In this paper, we presented state-of-the-art research carried out in the field of patient health monitoring with different technologies such as IoT, wearable devices and cloud. The motive of this work is identifying different solutions and frameworks for monitoring patient health status by sensing physiological parameters and other information of patient. It also presents what are the current systems and technologies which give effective and efficient solutions to the patient in the perspective of real-time information/alters to be communicated to caretakers or other belongings of patient. With this review, we conclude that there is an acute need of patient monitoring system to help or save patient life by giving real-time decisions and actions without delays.

References

1. Ahmed, S., Millat, S., Rahman, M. A., Alam, S. N., & Zishan, M. S. R. (2015). Wireless health monitoring system for patients. In*2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON—ECE)*, Dhaka (pp. 164–167).
2. www.elprocus.com/automatic-wireless-health-monitoring-system-circuit/.
3. Aminian, M., & Naji, H. R. (2013). A hospital healthcare monitoring system using wireless sensor networks. *Journal of Health & Medical Informatics, 4,* 121. https://doi.org/10.4172/2157-7420.1000121.
4. Ali, Z., Muhammad, G., & Alhamid, M. F. (2017). An automatic health monitoring system for patients suffering from voice complications in smart cities. *IEEE Access, 5,* 3900–3908.
5. World Health Organization (WHO). Are You ready? What You Need to Know about Ageing. Available online: http://www.who.int/world-health-day/2012/toolkit/background/en/. Accessed on 11 May 2017.
6. Kulik, C.T., Ryan, S., Harper, S., & George, G. (2014). Aging populations and management. *Academy of Management Journal*,*57*, 929–935.
7. Yamada, I., & Lopez, G. (2012). Wearable sensing systems for healthcare monitoring. In*Proceedings of the Symposium on VLSI Technology*, Honolulu, HI, USA. 12–14 June 2012 (pp. 5–10).
8. Baig, M., & Gholamhosseini, H. (2013). Smart health monitoring systems: An overview of design and modeling.*Journal of Medical Systems*,*37*, 1–14.
9. Sanghavi, J. (2019). Review of smart healthcare systems and applications for smart cities. In*ICCCE 2019* (vol. 570, pp. 325–331).
10. Deepasri, T., Gokulpriya, M., Arunkumar, G., Mohanraj, P., & Shenbagapriya, M. (2017). Automated paralysis patient health care monitoring system.*South Asian Journal of Engineering and Technology*,*3*(2), 85–92.

11. Gómez, J., Oviedo, B., & Zhuma, E. (2016). Patient monitoring system based on internet of things. In*The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016).*
12. Krishnan, D. S. R., Gupta, S. C., & Choudhury, T. (2018). An IoT based patient health monitoring system. In*2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, Paris (pp. 1–7).
13. Saha, H. N., et al. (2017). Health monitoring using internet of things (IoT). In*2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*, Bangkok (pp. 69–73).
14. Abdelgawad, A., Yelamarthi, K., & Khattab, A. (2016). IoT-based health monitoring system for active and assisted living. In*International Conference on Smart Objects and Technologies for Social Good (GOODTECHS)*, Venice, Italy.
15. Patel, S., Singh, N., & Pandya, S. (2016). IoT based smart hospital for secure healthcare system.*International Journal on Recent and Innovation Trends in Computing and Communication*,*5*(5), 404–408.
16. Kodali, R. K., & Mahesh, K. S. (2017). Smart emergency response system. In*Proceedings of IEEE Region Conference (TENCON)*, Penang, Malaysia, Nov. 2017 (pp. 712–717).
17. Dar, B. K., Shah, M. A., Islam, S. U., Maple, C., Mussadiq, S., & Khan, S. (2019). Delay-aware accident detection and response system using Fog computing. *IEEE Access, 7,* 70975–70985.
18. Wu, X., Dunne, R., Yu, Z., & Shi, W. (2017). STREMS: A smart real-time solution toward enhancing EMS pre-hospital quality. In*2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Philadelphia, PA (pp. 365–372).
19. Rahman, Md. S., & Avi, Md. T. R. (2015). Exoskeleton Arm: The first step of real life iron suit. PhD thesis.
20. Tejaswi, U., & Suresh Kumar, P. (2016). Diagnosing diabetes using data mining techniques.*International Journal of Scientific and Research Publications*,*7*(6), 705–709.
21. Rishika Reddy, A., & Suresh Kumar, P. (2016). Predictive big data analytics in healthcare. In*Proceedings of IEEE 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT)*, Ghaziabad (pp. 623–626).
22. Suresh Kumar, P., & Pranavi, S. (2017). Performance analysis of machine learning algorithms on diabetes dataset using big data analytics. In*Proceedings of IEEE 2017 International Conference on Infocom Technologies and Unmanned Systems (ICTUS'2017)*, Dubai, United Arab Emirates (UAE), December 2017 (pp 580–585).
23. Chaudhary, A., Peddoju, S. K., & Peddoju, S. K. (2020). Cloud based wireless infrastructure for health monitoring.*Virtual and Mobile Healthcare*, 34–55.
24. Peddoju, S. K., Kavitha, K., & Sharma, S. C. (2019). Big data analytics for childhood pneumonia monitoring. In*IGI Global* (pp. 1–17).
25. Peddoju, S. K., Upadhyay, H., & Bansali, S. (2019). Health monitoring with low power IoT devices using anomaly detection algorithm. In*Proceedings of IEEE Conference FMEC—2019*, Rome, Italy, June, 2019.
26. Peddoju, S. K., Upadhyay, H., & Lagos, L. (2020). File integrity monitoring tools: Issues, challenges, and solutions.*Concurrency and Computation: Practice and Experience*, e5825. https://doi.org/10.1002/cpe.5825.
27. Suresh Kumar, P., Ramachandram, S. (2019). Fuzzy-based integration of security and trust in distributed computing. In*Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing* (vol. 816). Singapore: Springer.
28. Peddoju, S. K., & Upadhyay, H. (2020). *Evaluation of IoT Data Visualization Tools and Techniques.* Data Visualization: Springer.
29. Peddoju, S. K., Upadhyay, H., Soni, J., & Prabakar, N. (2020). Natural language processing based anomalous system call sequences detection with virtual memory introspection.*International Journal of Advanced Computer Science and Applications (IJACSA)*,*11*(5). http://dx.doi.org/10.14569/IJACSA.2020.0110559.

Workflow Management Through Database Migration in Cloud Computing Environment

Amit Kumar, M. SivakKumar, and Varsha Namdeo

Abstract Cloud computing (CC) is a new concept compared with conventional methods to information technology, in which companies and enterprises implement open cloud systems on a pay-per-use basis to minimize costs. Virtualization is a pervasively implemented strategy in modern data centers (DCs) to optimize the utilization of energy, minimize greenhouse gas emissions, and decrease operating costs. The migration of virtual machines (VM) is commonly used within and through DCs to address a range of virtualized cloud system needs. For example, server consolidation needs power management via VM migration. Load balancing, fault tolerance, network repair, and reducing the violation rate of service-level agreement (SLA) often include live VM migration. The VM migration process is very resource-intensive and involves smart solutions to prevent saturation of the network bandwidth which reduces server downtime.

Keywords Cloud computation · Virtual machines · Data centers · Database migration · Linear regression

1 Introduction

Cloud computing is the idea of a "Computing Environment," under which services are deployed and the results of research are released in a regular Web browser window on a local computer, including all software and their data available for service on a remote Internet server [1]. The benefits of cloud computing include: decreased computational capacity demands for PCs, improved fault tolerance and reliability,

A. Kumar (✉) · M. SivakKumar · V. Namdeo
Department of Computer Science & Engineering, SRK University, Bhopal, India
e-mail: amitpccollege2008@gmail.com

M. SivakKumar
e-mail: Sivakum.slm09@gmail.com

V. Namdeo
e-mail: Varsha_namdeo@yahoo.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_36

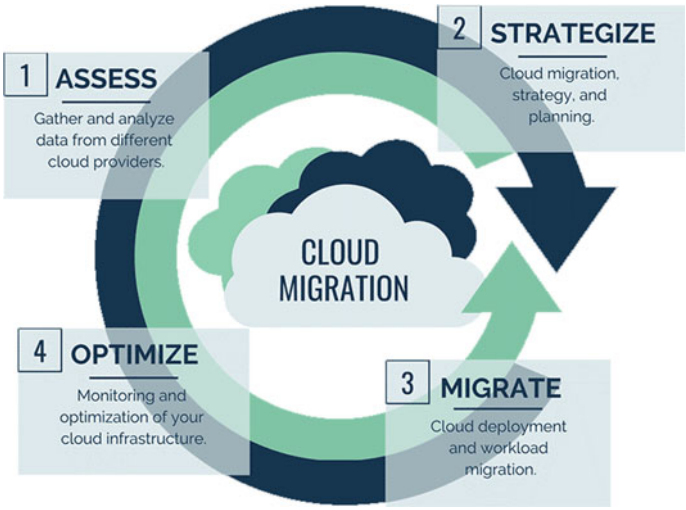


Fig. 1 Methodology for cloud migration

multiple data processing speeds increase, hardware and software costs, repair costs, power management as well as disk space savings.

This research focuses mainly on evaluating cloud computing performance and comparing different load balancing algorithms using the network simulator cloud analyst.

Workflow Management. Cloud computing may have many definitions. According to [2] "Cloud Computing is a kind of distributed and parallel network consisting of a series of interconnected as well as a virtualized form of computers that are dynamically supplied and displayed as one or more centralized computing resources dependent on service-level agreement amid service provider and customer." [3] claims that there are several different cloud types, such as public cloud, private cloud, group cloud, and hybrid cloud. Those are cloud deployment models.

Cloud Migration Framework. Migrating workloads into cloud models is characteristically an "application-driven" movement where each picture/case in the cloud commonly runs a singular application workload. All things considered, moving workloads to "cloud conditions" need to take after a multi-step procedure to get those applications running effectively in the focused on cloud condition. To start with, the focus on applications should be recognized and "segregated" from alternate applications running on that same server. At that point, a picture of that application, its fundamental operating system, and framework administration agents should be made and added to the cloud index [7]. At last, the picture should be instantiated in the cloud condition and confirmed to keep running with adequate quality of service (QoS) attributes. The specialized contemplations for migration can be condensed as:

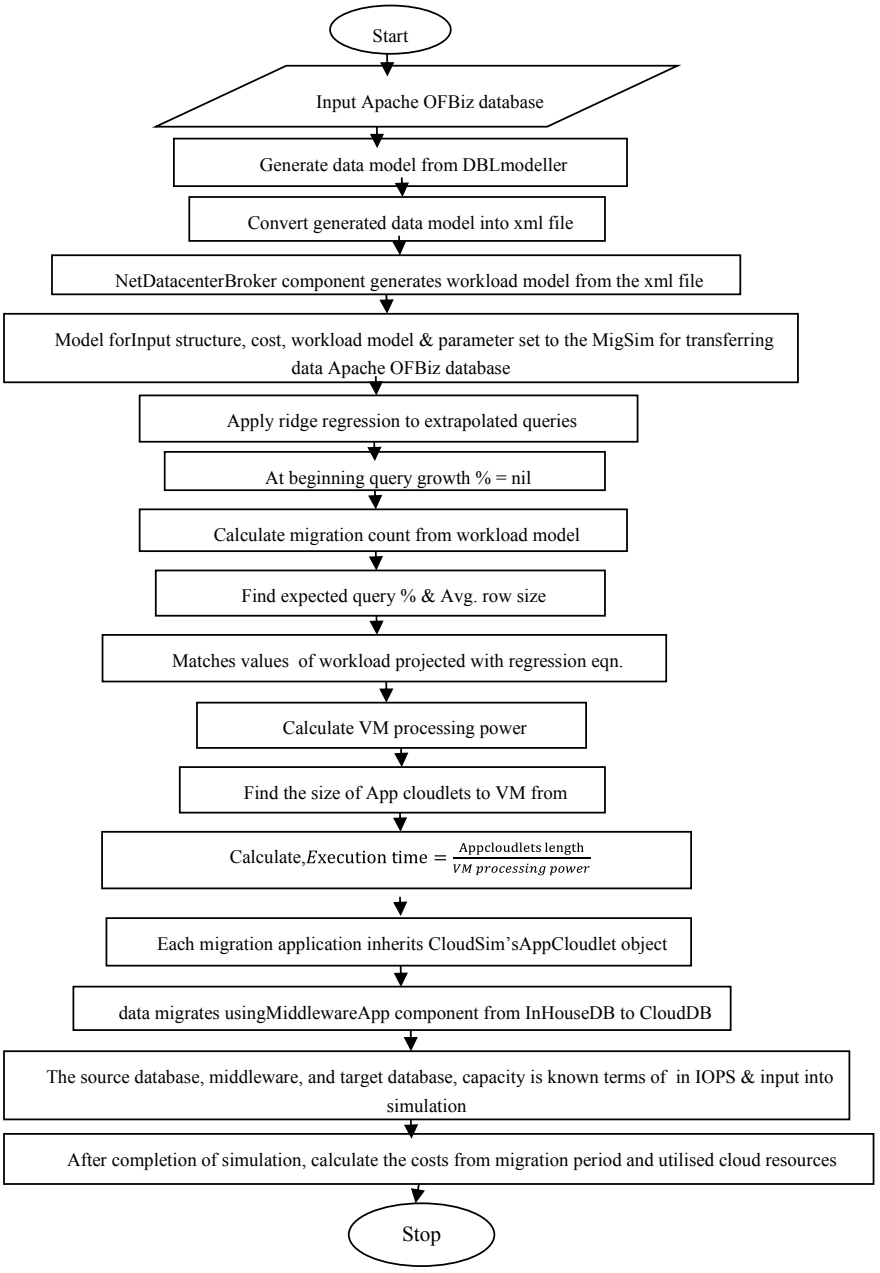


Fig. 2 Flow diagram of the proposed methodology

Fig. 3 Total time taken by MigSim simulation

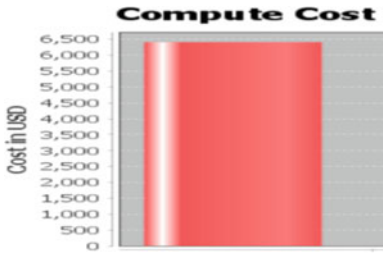


Fig. 4 Computation cost taken by MigSim simulation

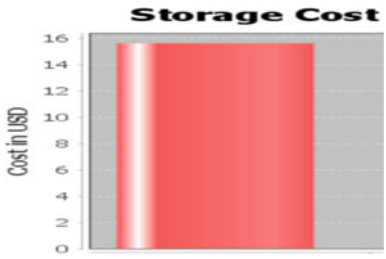


Fig. 5 Storage cost taken by MigSim simulation

- Software compatibility
- Reference architecture
- Workload characteristics
- Platform needs over the majority of the migration patterns, and there are some regular themes.

Database Migration Across Clouds. The time it takes to finish the real migration of the objects and data from one type of database is less in comparison with the amount of time to finish a general migration from evaluation to generation rollout. Migration of a social database to others is relatively less demanding than migrations of a non-social database to a social database because the association of items in a social database is very comparative contrasted with non-social databases, for example, databases of hierarchical and network type. All social database merchants likewise offer apparatuses that give vigorous relocation abilities in a computerized form. Despite the level of mechanization and achievement factor of any migration apparatus, be that as it may, here and there manual mediation will be required while

moving beginning with one database then onto the following. Migration of database tasks can be divided into given categories:

- Migration of database schema
- Migration of data
- Migration of database stored program
- Application migration
- Database administration script migration
- The capability to manage in-house databases is insufficient
- IT is not a central functional unit
- You are an SME and need to cut beginning capex cost
- You are working with novel applications or creating one, and need to attempt the cloud as a testing situation
- Moving to the cloud for your disaster recovery backup and utilizing it as a preliminary run to recognize issues and hurdle to dataset migration.

2 Data Migration Strategy Realization

Here, initially, the entire stored data is within a private cloud. There are two kinds of migration: from private cloud to public, then again reverse to the private cloud. We are primarily concerned regarding the former, moreover are proposing a self-adaptive periodic migration approach for part-migration retention [5]. The implementation ring protocol including detailed questions is as follows.

Which objects require migration? As already stated, it is very important to schedule which data to move. Popular data transfer algorithms such as (1) Store space solely [14]: while using strategy, private cloud objects are only migrated at the public depending on available storage space. (2) FIFO: Here, files previously maintained would be transferred to the Internet, irrespective of the level of use. (3) LRU: Handles every file sizes on an equal basis, but the size of the file dictates the cost of the migration. (4) Dimension only: Big items are continuously moved but only used.

Such migration algorithms solve the problem to some degree. But integrating them, which means considering exhaustion of storage capacity, amount of use of objects, and at the same time, the scale of objects would produce a better outcome. We refer to circumstances like “factors of migration.” First, we have to work out their particular definition, and how it is measured.

2.1 The Definition and Migration Factors Calculation

2.1.1 Definition

Space Saturation. $\ominus = S_{\text{used}}/S_{\text{total}}$. Assume$S_{\text{used}}$ denote storage space size and its usage in the private cloud, andS_{total} signifies the whole size of storage pf private cloud. Thus, $\ominus$ stands for cloud storage pf private form space saturation. The $\ominus$ andU_{start} comparison that is a situation to find out the time of migration. Once $\ominus <$ U_{end}, the process of migration ends.U_{start} andU_{end} are later being defined.

ObjectSize. Object size specifies a single object size of byte as a data unit. Assume size consider note of object size threshold that to migrated.

ObjectSize $\geq$ Size

LastUseTime. The aspect indicates one object's last time of use. As defined below, the to-be-migrated object's last-use time threshold is Week. Day = 14, for ex: means that objects were not made use in the previous 2 weeks would be migrated to BOS.

(DateOfToday-LastUseTime) $\geq$ Day

Space Saturation. $\ominus = S_{\text{used}}/S_{\text{total}}$. Assume$S_{\text{used}}$ denotes storage space size and its usage in the private cloud, andS_{total} signifies the whole size of storage pf private cloud. Thus, $\ominus$ stands for cloud storage pf private form space saturation. The $\ominus$ andU_{start} comparison is a situation to find out the time of migration. Once $\ominus < U_{\text{end}}$, the process of migration ends.U_{start} and U_{end} are later being defined.

2.1.2 Calculation

Saturation of space is much easy to get between the three migration factors. But how does ObjectSize and LastUseTime get? Here, we use to put forward a single solution: to retain a new database of MySQL with activity S3. Swift3 middleware supports system, bucket, and task processes. We center primarily on the object service, consisting of object GET, object DELETE, object PUT, and object HEAD. We record the objects' current state in the database when running certain operations.

2.2 Process

2.2.1 When Migration Performed?

The strategy of data migration has to be efficient and non-intrusive to maximize system resource usage. To boost performance, the collection of suitable migration factors is as described. Non-intrusive is where the process of migration needs to reduce its effect on parallel execution of workloads for storage [8]. And the time to conduct migration counts a lot. They suggest a self-adaptive periodic migration approach that is part-migration-part-retention. They ought to introduce two words:

Periodicity. Periodicity is performing replication of data at the period defined by the program. For e.g., if the data collected regularly is not many, we might opt to switch once a while in a week access control frequency is smaller, such as 22:00–6:00. Unix such as OS can use Cron functionality to strategize the process of migration.

Self-Adaption. Once it is a moment of migration, now do not move all the artifacts which match the condition of migration from the Swift to the BOS, other than determining ardently whether or not to migrate depending on current saturation of the storage room. The following factors must be put away: Ustart and Uend. Ustart represents the exhaustion of the fast storage space, causing migration of data. Migration starts when $\ominus$ is higher than Ustart. Uend is the point of exhaustion of privately reserved space that stops the migration. Such two attributes are user-configurable based on cloud computing system use. This will, on the one hand, protect Swift storage from an explosion in space. In comparison, migration does not occur when the space consumption ratio is small.

2.2.2 Execute the Process of Migration

The procedure in detail of second also is given:

Swift Operation. Steps 5th, 9th, 11th, and 14th include Swift's operations: join to Swift Web site, access object locally, then remove it from the Swift and also the central server. Python is an apt language to execute the migration, since it is being written in Swift. On Swift data, two methods can execute operations named CRUD: the interface boto's S3 as well as the Python cloud files.

BOS Operation. Phase 13 is to link to the BOS, and add local objects to it. The object that is transferred to BOS is called according to a definite layout to prevent redundant names:-*account_container_object*.

MySQL Operation. The task parameters displayed in Step 1st are limits, the scale, and the day. For e.g., the administrator may opt to move items that were not visited in three months with a size greater than 100 M, by locating Day $=$ 90, size $=$ 100*1024*1024 $=$ 104857600B.subsequent to migration finished and update the Table ObjectInfo**updateDatabase()**. The *Location*record is being updated starting from "Local" to the "BOS" soon after migration of better type.

2.2.3 Migration Log

The migration method is separated to: (1) connecting to Swift server; (2) downloading objects via Swift; (3) storing objects locally; (4) uploading objects from the local cache to BOS; (5) deleting objects as of local cache; (6) extracting objects from Swift.

Since intermittent migration also takes place when using the device infrequently, often nighttime, we need not rule out the probability of exception, such as a network link malfunction and account verification problems. Except for (5), the migration for

the object is over after an error occurs in the process. It is therefore important to note the progress of the migration, time to complete each step, and explanation of why such exceptions occur.

2.3 Performance

We carried out the migration algorithm's total performance check (on a 1 GB memory of VM, two processors, hard disk OF 100 GB SCSI using Ubuntu server enabled, with Swift's SAIO setup having several proxy server 1 and storage nodes are 4). The time of migration of an object is set as:

$$T_M = T_D + T_U + T_0$$

where:$T_D$is given time to download items from Swift storage Server to a local cache.T_0: It is time to add a local cache item to the BOS. Another time usage in object relocation, such as eliminating local cache function and Swift, upgrading MySQL client operations.T_M: It is directly correlated with the object size to be migrated. Sometimes, the PUT process with BOS accounts; nevertheless, it is connected to the bandwidth of real-time, device memory, etc.

Challenges as well as show how techniques for migrating data to a NoSQL cloud database are utilized.

3 Research Methodology

In this work, data have collected from the Apache OFBiz which is Oracle and MySQL type in the domain of Business Management and E-commerce. We have finished analyzing facets of our methodology. Firstly, a database is extracted from "DBLModeller" database models, then we address and analyze the accuracy of the cost and length of database migration estimates created by the method, and their estimate of cloud service operating costs in "Migration" as well as "Running costs," respectively. We made use of linear regression, precisely the ordinary least squares (OLS) approach [13], to predict the number of possible queries to read and write the database receives.

Linear Regression: A linear equation is used to model relationships between two variables by implementing a linear regression to observational data. An illustrative variable can and will be considered as dependent on another variable. For instance, a simulation firm may want a linear regression model to appropriate labels weights to their levels [14].

Input. Apache OFBiz database (Business Management & E-commerce).

Step 1: DBLmodeller generates a data model. These data are in the form of tables.

Step 2: Now, check the records and convert the generated data model into an xml file.
Step 3: The NetDatacenterBroker component generates a workload model from the xml file for the InHouseDB. Initially, input the structure and cost model, workload model, as well as a parameter set to the MigSim for transferring data. Cost reads from the existing database.
Step 4: We have used ridge regression to calculate approximately future read and write queries number those received via the database.

(i) Initially query growth percentage set to 'nil.'
(ii) Calculate migration count from workload model
(iii) Find the expected query growth percentage by ridge regression and also find the average row size of the tables

Step 5: Workload values estimated with the regression equation is matched.
Step 6: Find VM processing power by
*No. of processor * MIPS (millions of instructions per seconds) + Bandwidth*
Step 7: Find the size of Appcloudlets to VM from reading/write queries.
Step 8: Then, update the virtual machine allocation policy to scheduling the cloudlets
Step 9: Either each database or the application migration inherits the AppCloudlet object of CloudSim's.
Every AppCloudlet is being assigned to a VM, running on top of a detached physical host in datacenter of CloudSim.
Step 10: The component name MiddlewareApp migrates data from InHouseDB to CloudDB via slack capacity.
Step 11: The source database, middleware, as well as target database capabilities are specified in IOPS and most are input further in simulation. IOPS (termed as or input or as output/second) is metric of storage devices performance such as SSDs and services of cloud storage.
Step 12: After completion of the simulation, calculate the costs from the duration of migration and cloud resources that are utilized.

4 Result and Discussion

Three pieces of information that are crucial during database migration to the cloud are the duration, future running, and migration cost. Result visualization of the MigSim based on time, cost, and storage cost is given in figures and Table 1.

Table 1 Parameter representation of Migsim

Parameters	MigSim
Time	1,000,000 + ms
Computation cost	6250 USD
Storage cost	15.5 USD

5 Conclusion

To prevent problems with data migration, the organization requires a secure approach to allow data migration to be planed, validated, and migrated. With the rigorous standards and procedures, there is a need to speed up and migrate data to the cloud with automatic transfer of data and convergence through control cost and capital for disaster recovery. Data scalability is a major concern for cloud providers when cloud data is being stored or migrated. Cloud services such as Amazon, Google Cloud, and Microsoft Azure are being used by businesses to reduce big data problems and their migration. Cloud computation cost, storage cost, and time are the main parameters of the accurate measurement of the simulation tool.

The future work aims at developing Meta legacy system as well as cloud application models, which will become an abstract model for defining Meta model transformation, to explore the benefits of model paradigms.

References

1. Duan, Zhiying, & Cao, Yizhen. (2014). Serv-U based audio data transmitting management system on broadcasting stations. *Applied Mechanics and Materials, 543–547,* 2973–2976.
2. Kuo, Y. H., Jeng, Y. L., Chen, J. N. (2013). A hybrid cloud storage architecture for service operational high availability. In*IEEE 37th Annual Computer Software and Applications Conference Workshops*, Japan (pp. 487–492).
3. Grewal, R. K., & Pateriya, P. K. (2013). A rule-based approach for effective resource provisioning in a hybrid cloud environment. *Advances in Intelligent Systems and Computing,* 41–57.
4. Patel, M., Patel, R., & Chaube, A. R. (2015). Hybrid cloud computing: Data sharing and security issues. *International Journal of Research, 2*(1), 464–467.
5. Breiter, G., Naik, V. K. (2013). A framework for controlling and managing hybrid cloud service integration. In*IEEE International Conference on Cloud Engineering,* Redwood City, CA (pp. 217–224).
6. Voras, I., et al. (2011). Evaluating open-source cloud computing solutions. In*Proceedings of the 34th International Convention for Information and Communication Technology,* Opatija (pp. 209–214).
7. Yadav, S. (2013). Comparative study on open source software for cloud computing platform: Eucalyptus, OpenStack and opennebula. *Research Inventory: International Journal Of Engineering And Science, 3*(10), 51–54.
8. Barkat, A., Dos Santos, A.D., & Ho, T. T. N. (2014). OpenStack and CloudStack: Open source solutions for building public and private clouds. In*2014 16th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, Timisoara (pp. 429–436).

9. Iqbal, A., & Colomo-Palacios, R. Key opportunities and challenges of data migration in cloud: Results from a multivocal literature review. In*CENTERS-International Conference on ENTERprise Information Systems/ ProjMAN - International Conference on Project MANagement/ HCist - International Conference on Health and Social Care Information Systems and Technologies.*
10. Tomar, D., Bhati, J. P., Tomar, P., & Kaur, G. (2019). Migration of healthcare relational database to NoSQL cloud database for healthcare analytics and management. *Healthcare Data Analytics and Management Advances in Ubiquitous Sensing Applications for Healthcare*, 59–87.
11. Mao, B., Yang, Y., Wua, S., Jiang, H., & Li, K. -C. (2019). IOFollow: Improving the performance of VM live storage migration with IO following in the cloud.*Future Generation Computer Systems,91*, 167–176.
12. Alia, H., Moawadb, R., & Hosnia, A. A. F. (2016). A cloud interoperability broker (CIB) for data migration in SaaS.*Future Computing and Informatics Journal,* 1(1–2), 27–34, (2016).
13. Xiao, H., Hu, Z., & Li, K. (2019). Multi-objective VM consolidation based on thresholds and ant colony system in cloud computing.*IEEE Transactions.*
14. Zhou, X., Li, K., Liu, C., & Li, K. (2019). An experience-based scheme for energy-SLA balance in cloud data centers.*IEEE Transactions.*
15. Huang Dongmei, D., & Yanling, H. Q.. (2014). Migration algorithm for big marine data in hybrid cloud storage. *Journal of Computer Research and Development, 51*(1), 199–205.

A Pixel Count Approach for Lossy Image Compression

Tanuja R. Patil and Vishwanath P. Baligar

Abstract Digital image storage and transmission plays a very important role in today's modern world, as most of the data transfer involves images. Hence, digital image compression is of great importance. The compression leads to either lossy or lossless type of images. Here, we discuss about a unique approach for lossy image compression, which involves a threshold. The number of pixels whose sum is lesser than a threshold is counted, and this count is saved in a file instead of actual pixel intensity values. A difference between the threshold and the computed sum is also stored. Later, reconstruction is done by reading the count and difference values. Average is calculated, and the count number of pixels is replaced with this value. We find that we can achieve better quality at lower PSNR values with this approach as compared to JPEG algorithm.

Keywords Low PSNR · Lossy image compression · Threshold · Comparison with JPEG

1 Introduction

Digital images are inevitable in the transfer of information nowadays. They require large amount of memory for storage as well as for transmission, and also, time consumed to transmit is very high. Hence, it is very much essential that images are to be compressed. Some applications require lossless compression such as medical imaging, satellite imagery, but some applications like multimedia and GIS prefer highly compressed data rather than high-quality images. Based on these needs, we have lossless and lossy type of image compression techniques.

T. R. Patil (✉) · V. P. Baligar
K.L.E. Technological University, Hubballi, India
e-mail: tanuja_p@kletech.ac.in

V. P. Baligar
e-mail: vpbaligar@kletech.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_37

Compression is possible because of redundancy that exists in the digital image storage data. It may be in the form of coding, inter-pixel, and psycho-visual redundancy. By using variable length coding and some transform techniques, we can reduce these redundancies [1, 2].

In lossless compression, it is very much essential that the image has to be reconstructed accurately to achieve high quality. Lossless approaches use techniques like Huffman coding, LZW coding, run length coding, and arithmetic coding. [3]

Here, we present a lossy approach for image compression over grayscale images. Section 2 is literature survey, which provides an overview about some image compression techniques. In Sect. 3, we discuss about our proposed algorithm. Section 4 explains the compression and decompression algorithms. In Sect. 5, we discuss about 'correctness ratio,' and in Sect. 6, we provide experimental results, and Sect. 7 gives a comparison between our algorithm and the JPEG algorithm. Section 8 gives the conclusion.

2 Review of Literature

In this section, we discuss about some of the lossy and lossless image compression techniques and their outcomes.

As per Weinberger et al., 2000, low complexity lossless compression for images (LOCO-I) is the algorithm at the core of the new ISO/ITU standard, JPEG-LS for continuous-tone image compression. This algorithm gives good compression ratio, and level of complexity is also less [5].

Raid et al. discuss about lossy image compression algorithm using DCT which is used for full-color still image applications [6].

Baligar et al., 2006, discuss about the image coding algorithm based on fixed threshold method. Threshold is the peak absolute error (PAE) allowed in the decompressed image. Here, a comparison is made with SPIHT algorithm to show that this algorithm gives visually better images and execution time is less [7].

As per Sinisa ILIC, Mile PETROVIC, Branimir JAKSIC, Petar SPALEVIC, at lower values of bit rate, there arises noise effects from the compression methodology used in JPEG. Here contour-like structures appear, which are uncomfortable for better visibility [8].

Patil et al., discuss about a lossy compression algorithm using surrounding pixels method. Here, it is shown that at low PSNR levels, the number of exact pixels in reconstructed image increases, thus reducing the contour effects that may arise in JPEG at same PSNR values [9].

As per [4, 10–12], lossless techniques are discussed.

From the literature review, we understand that standard JPEG has some adverse effects at low PSNR values. Hence, we propose a pixel count approach by which quality can be improved at low PSNR values.

3 Pixel Count Approach Using Threshold Method

In this section, we present a pixel count approach using threshold algorithm. A grayscale image is processed in raster scan manner. Here, number of pixels is computed whose sum is lesser than a threshold, and this number is stored in a file instead of storing each pixel intensity value. Later, reconstruction is done by calculating the average of this count in a unique way, and each pixel value is thus reconstructed.

Here, threshold is taken as 255, and sum is computed as

$$\text{sum} = f[x, y] + f[x, y + 1] + \ldots <= 255$$

Later, difference is calculated as

$$\text{Difference} = 255 - \text{sum}$$

This count is saved in a file, and difference is saved in another file.
Reconstruction is done as shown below.
An average value is computed using

$$\text{Average} = \frac{255 - \text{difference}}{\text{Count}}$$

The number of pixels (=count) is replaced by this average value which is near to the actual value.

4 Methodology

Here, we describe the algorithms with examples taken over a sample Lena image. For our work, we have used the standard set of grayscale images of size 512 × 512.

4.1 Algorithm Used for Compression

1. Input the grayscale image pixel intensity values.
2. In the raster scan manner, count the number of pixels whose sum <=255

$$\text{Sum} = f[x, y] + f\,[x, y + 1] + \ldots f[x, y + n] \tag{1}$$

Condition is checked using Eq. (2)

$$f[x, y] + f[x, y+1] + \ldots f[x, y+n] \,(\text{count of pixels}) <= 255(\text{threshold}) \tag{2}$$

3. Store this count in a 'count' file.
4. Find the difference using Eq. (3)

$$\text{Diff} = \text{sum} - 255 \tag{3}$$

5. Store this difference in 'difference' file
6. Huffman encode the files.

Outputs generated are count value file and difference file.

4.2 Algorithm Used for Decompression

We can reconstruct the image using following algorithm using 'count' and 'difference' as input files

1. Input the values from count and difference files generated by compression.
2. Do the Huffman decoding.
3. Declare an array for image reconstruction.
4. C = count value, D = difference value
5. Diff = 255 − D
6. Reconstruct 'C' number of pixels with a value 'Avg'
   where Avg = Diff/C

Examples
We have applied the pixel count approach using threshold over grayscale images. Here, we discuss this algorithm with examples of two sample sets of Lena image.

Compression:

The input image is scanned in raster scan manner. The pixel intensity values of adjacent pixels are added till it is lesser than a threshold. In this case, threshold is taken as 255. The count value is stored in a file say 'count.' The difference between the sum and threshold is calculated and stored in a file say 'difference.' Figure 1a, b show two sample sets of Lena image.

Calculations for sample set 1:

1. Initially, first pixel which is 162 < 255 is considered (Fig. 1a). Since it is lesser than the threshold, i.e., 255, next pixel value is added which is again 162.
   162 + 162 = 324 which is greater than threshold, 255. Hence, count is stopped at 1 and the count is saved as '1' in 'count' file.
2. Next, compute the difference as, Difference = 255 − 162 = 93
   Ninety-three is stored in 'difference' file.

162	162	162	161	162	157
163	164	164	157	158	161
159	159	160	160	158	155

(a)

106	110	108	111	112	108
101	104	104	107	113	111
102	99	105	108	114	114

(b)

Fig. 1 **a** Sample set 1 of Lena image. **b** Sample set 2 of Lena image

Calculations for sample set 2:

1. In second sample (Fig. 1b), first pixel value is 106 and seccond pixel value is 110.
   $106 + 110 = 216$. If we add next pixel value, i.e., 108, $216 + 108 = 324$ which is greater than threshold. Hence, count is '2', and this '2' is saved in 'count' file.
2. Now compute the difference as $255 - 216 = 39$ and '39' is saved in 'difference' file.

Similarly, count and difference are computed for next pixels and stored in 'count' file as shown in Fig. 2a, b and 'difference' file as shown in Fig. 3a, b.

Decompression

Calculations for reconstruction:

Initially, encoded 'count' and 'difference' files are read and Huffman decoded. A two-dimensional array arr[i, j] is declared to store reconstructed values.

A 'count' value is read as 'C', and 'difference' value is read as 'D'.

1. For sample set 1, first values from the two files are $C = 1$ and $D = 93$

1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	1	1

(a)

2	2	2
2	2	2
2	2	2

(b)

Fig. 2 **a** Count file for sample set 1. **b** Count file for sample set 2

93	93	93	94	93	98
92	91	91	98	97	94
96	96	95	95	97	100

(a)

39	36	35
50	44	31
54	42	27

(b)

Fig. 3 **a** Difference file for sample set 1. **b** Difference file for sample set 2

Reconstruction is done using the equation, Avg = (255 − *D*)/*C*, Avg = (255 − 93)/1 = 162. Here, the count value '*C*' is 1. Hence, Avg value is assigned to first pixel which is exactly same as original, i.e., 162.

2. For sample set 2, *C* = 2 and Avg = (255 − *D*)/*C*, Avg = (255−39)/2 = 108.

Here, first two-pixel values are assigned the value 108, instead of 106 and 110 as 'Avg' value is assigned to '*C*' number of pixels where it results in slight loss.

Though it is lossy, reconstructed image quality is good which is measured by a metric called 'correctness ratio,' which is explained in Sect. 5. For the sample set 1, reconstruction is exactly same as original, but for sample set 2, there is slight loss. By this algorithm, we get exact reconstruction wherever the pixel intensity values are higher and results in no compression. But, we get more count wherever the image is darker and results in more compression ratio. (This is clear by the size of count and difference files for sample set 2).

5 Correctness Ratio—A New Quality Metric

The quality of reconstructed images can be assessed using metrics MSE and PSNR. But, here we propose a metric called correctness ratio (Co.R.) which gives a better method to test the quality of reconstructed images [8].

$$\text{Co.R.} = \frac{\text{Total of actual pixels in reconstructed image as that of original}}{\text{Total pixels in the original image}}$$

This threshold approach gives more accuracy as compared with JPEG, at same PSNR values, which is shown in Sect. 7. The number of actual pixels in the reconstructed image increases, thus increasing the correctness ratio. As this ratio increases, contours which may appear after reconstruction are reduced.

6 Results

We got the following results when we applied the proposed algorithm on standard set of images as shown in Fig. 4. Here, the reconstructed images show the quality of the images which seem to be near to original. Left side is the original images, and right side is the reconstructed images.

Fig. 4 Standard set of images with reconstructed images

Table 1 Comparison of number of correct pixels as that of original

Input files	Increased number of pixels as per proposed algorithm, compared to JPEG	Co.R. of JPEG	Co.R. of proposed approach
Lena	49,775	0.18	0.32
Baboon	72,867	0.05	0.33
Barbara	37,681	0.05	0.19
Airplane	75,764	0.1	0.38
Aya_matsuura	54,697	0.09	0.3
Pepper	45,656	0.08	0.23

7 Comparison Results

This section gives a comparison between the threshold algorithm with that of JPEG lossy. We compute the number of pixels in the reconstructed image, which are exactly same as that of original image using both algorithms, and we found that it is increased with proposed approach as shown in Table 1. Similarly, we compute correctness ratio for both algorithms, and it is found to increase as shown in Table 2.

8 Conclusion

The pixel count approach using threshold algorithm is an innovative and low computation-intensive method for image compression. The performance metric used, 'correctness ratio,' gives a count of, how many pixels have the same values as that of original. By this, we find that there is an increase in this number as compared to JPEG lossy, at same PSNR values. This shows that the contours which were appearing with JPEG algorithm can be reduced. Hence, we can say that the quality of reconstructed

Table 2 Comparison of correctness ratio (Co.R.)

Input files	PSNR in dB	Total pixels in the original image	Correct pixel count using JPEG approach	Correct pixel count using proposed approach
Lena	34.15	262,144	34,454	84,229
Baboon	31.2	262,144	14,322	87,189
Barbara	31.69	262,144	13,470	51,151
Airplane	34.5	262,144	26,336	102,100
Aya_matsuura	32.5	262,144	24,132	78,829
pepper	31.1	262,144	16,580	62,236

image is improved with this approach and compression ratio achieved is around three for standard data sets. Further, it can be tested for different threshold values and adaptive techniques for improvement.

References

1. Kaur, R., & Choudhary, P. (2016). A review of image compression techniques. *International Journal of Computer Applications*, *142*(1) (0975-8887).
2. Anju, & Ahlawat, A. (2016). Performance analysis of image compression technique. *International Journal of Recent Research Aspects*, *3*(2). ISSN 2349-7688.
3. Gonzalez, R. C., & Woods, R. E. (1978). *Digital image processing* (2nd ed.). Pearson Prentice.
4. Baligar, V. P., Patnaik, L. M., Nagabhushan, G. R. (2003). High compression and low order linear predictor for lossless coding of grayscale images. *Image & Vision Computing*, *21*, 543–550. www.elsevier.com.
5. Weinberger, M. J., Seroussi, G., & Sapiro, G. (1996). LOCO-I: A low complexity, context-based, lossless image compression algorithm. In *Proceedings of 1996 Data Compression Conference*, Snowbird, UT, Mar. 1996 (pp. 140–149).
6. Raid, A. M., Khedr, W. M., El-dosuky, M. A. & Ahmed, W. (2014). JPEG image compression using discrete cosine transform-A survey. *International Journal of Computer Science & Engineering Survey (IJCSES)*, *5*(2).
7. Baligar, V.P., Patnaik, L.M., Nagabhushan, G.R. (2006). Low complexity and high fidelity image compression using fixed threshold method. *Information Sciences, 176*, 664–675.
8. Ilic, S., Petrovic, M., Jaksic, B., Spalevic, P., Lazic, L., Milosevic, M. (2013). Experimental analysis of picture quality after compression by different methods. *Przegląd Elektrotechniczny*. ISSN 0033-2097, R. 89 NR 11/2013.
9. Patil, T. R., Baligar, V. P., & Huilgol, R. P. (2018). Low PSNR high fidelity image compression using surrounding pixels. In *2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)*, Kottayam, India (pp. 1–6). https://doi.org/10.1109/iccsdet.2018.8821082.
10. Novikov, D., Egorov, N., & Gilmutdinov, M. (2016). Local-adaptive blocks-based predictor for lossless image compression. In *2016 XV International Symposium "Problems of Redundancy in Information & Control Systems"*.

11. Huilgol, R. P., Baligar, V. P., & Patil, T. R. (2018). Lossless image compression using seed number and JPEG-LS prediction technique. In *Conference proceedings—Punecon 2018*.
12. Huilgol, R. P., Baligar, V. P., & Patil, T. R. (2018). Lossless image compression using proposed equations and JPEG-LS prediction technique. In *2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)*, Kottayam, India (pp. 1–6). https://doi.org/10.1109/iccsdet.2018.8821065.

Firmware Attack Detection on IoT Devices Using Deep Binary Pattern Classification Mining (FA-PCM)

E. Arul and A. Punidha

Abstract Because IoT devices began to invade our everyday lives, businesses would eventually have to tackle thousands of IoT gadgets, if not a large amount of them. There is a shortage of protection functionality where the device customers are limited or are unwilling to change the security tests. Deficient design of authentication is apparent when the device user interface does not allow choices for granular consumer approvals or where the usage of secure passwords is convincing. The risk here is that the IoT gadget can be made more easy to attack so that the gadget is unapproved. An algorithm for image processing is proposed in this paper Deep Binary Pattern Classification Mining (FA-PCM) for classifying the firmware assault in IoT smartphones.The FA-PCM classifies the IoT application files as a binary image of the proposed firmware detection process. The values extracted from the file values are converted into binary pixel image pattern scaling and deep pattern mining is applied. The inter and intra cluster matching of the binary pattern was determined. The continuous probability distribution for a real-valued random variable, such as lopsidedness and Leptokurtic, is measured and categorised by the FA-PCM deep mining of the various clustered pixels. The (FA-PCM) area with the highest value is classified as a malware server. The identification efficiency of the proposed system has been tested using 960 compact executable data. The 96.12% true positive identification form with 0.09% false positive. Preliminary results show that the new algorithm is better than other standard malware detection methods.

Keywords IoT · Information security · Malware · API calls · lopsidedness and Leptokurtic · Binary pattern classification · Gadgets attack

E. Arul (✉)
Department of Information Technology, Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India
e-mail: arulcitit@gmail.com

A. Punidha
Department of Computer Science and Engineering, Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India
e-mail: punitulip@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_38

1 Introduction

The Internet of Things (IoT) presents various advantages to buyers and can possibly change the manners in which that shoppers interface with innovation in key manners. Later on, the Internet of Things is probably going to merge the virtual and physical universes together in manners that are as of now hard to fathom [1]. An anticipated unavoidable presentation of sensors and gadgets into right now cozy spaces—for example, the home, the vehicle, and with wearables and ingestible, even the body—presents specific difficulties. As physical articles in our regular day-to-day existences progressively identify and share perceptions about us, shoppers will probably keep on needing protection [2].

The key issue is the privacy concerns linked to a mobile app with smart devices that require a client to comply with the device but that could cause an attacker to enhance unauthorized access to the gadget [3]. Unsafe web interface authentication vulnerabilities that could cause this problem include: database aggregation, powerless preferred reputation, network accreditation, database injection, session control, powerless password locking [4].

Lack of decryption/authorization—this jurisdiction handles unsatisfactory frameworks to search for IoT UI or possibly faulty approval elements, allowing a consumer to raise higher access thresholds at any time [5]. Overt security risks can contribute to: lack of password sophistication, inadequate protection of the passwords, lack of protection for independent variables, unreliable user retention, advantage increases, and the lack of function-dependent network access.

2 Related Work

Jiawei Su's study of computing research is a conventional cloud expansion that supports many connected technologies such as household appliances, network sensors, networking, and user interface controls [6]. This IoT is a typical media extensions. IoT implementations are more like microcomputers like conventional integrated domain computers [7]. Thus, a variety of new kinds of threats, usually affecting Internet and IoT computers, are illustrated. DDoS attacks have lately been shown in IoT systems because they neglect simple safety and security frameworks [5], and incidents such as the Mirai and protective measure botnets have lately been seen.

Rakesh Singh Kunwar's study of the greatest breakthrough in the digital world is social networking. It is the main site for documents, files and papers to be shared. He also expresses opinions, suggestions, and emotions using various methods and techniques [2]. We are hyper-connected and constantly exchange knowledge. To attackers, it is simple, and one of the simplest and more interesting features to execute vulnerabilities in such a situation is by JPEG pictures. Through the usage of obfuscation techniques technologies, perpetrators implanted legal or innocent photographs

with harmful codes [8]. This content is only some few pieces of keys which benefit from technology weakness. It allows the offender to navigate the justice procedure online. Our primary objective in this context is to recognize any codes or data in the picture. The central portion of this system is then focused on application discovery and its harmful implications. In this context, a workaround is given for the inclusion of malicious activities in JPEG photos on online social sites [9].

Xin Zhou has done research during this study, integrating target identification and the prevention of threats to computer intelligence. Through using the machine learning target detection program [10], the researcher was able to recognise the vulnerabilities. The ransomware has been viewed as a gray image, and the texture properties of the Gabor filter stripped in our research. The author used them focused on artificial intelligence in order to test the recognition of strongly random trees and tenfold cross-validations. With a 96.19% precision and a 97.51% alert level for malicious based on two different closest, and others the most nearby to K, test data are reasonably decent.

3 Theoretical Background

3.1 Delineation Deep Binary Pattern Classification Mining (FA-PCM) to Detect Firmware Attack on IoT Devices

The category of firmware for IoT device threats is shown in Fig. 1. Initially, gray photographs from IoT executable files are extracted in this framework. The IoT operating framework is linear by nature. The positive note is translated from 16 consecutive IoT runtime bits to decimal value [3]. The ported decimal value corresponds to the IoT gray image's various binary pattern in-density. The whole IoT executable picture binary bits are converted to gray picture. Grayscale photographs are observed by deep binary pattern classification mining (FA-PCM) and lopsidedness and Leptokurtic (I&L). The I&L are imbalance and impulse diffusion governs the executable IoT application's lopsidedness and Leptokurtic. In the context of symmetric representations, the stability or absence of symmetry was calculated at almost zero. The γ_3 is denoting the random variable x of lopsidedness and Leptokurtic in Eq. (1) [8, 11].

$$\begin{aligned}
E_x &= C_x(\alpha_z - h_c h_x \sin\theta) \\
&-(C_{Ry} - I_C)(h_s + w_a \cos\theta)h_p \sin\theta \\
E_b &= C_{Gy}\left(C_n h_b \sin\theta + \alpha_x \cos\theta + \alpha_b\right) \\
E_c &= C_B\left(-C_n h_c \cos\theta + \alpha_x \sin\theta + h_x h_b\right) \\
&-(C_w - I_{Gc})(w_c + h_x \cos\theta)h_x
\end{aligned}$$

$$\begin{aligned}
E : \rho\left(\frac{\partial u_r}{\partial t} + u_r\frac{\partial u_r}{\partial r} + \frac{u_\phi}{r}\frac{\partial u_r}{\partial \phi} + u_z\frac{\partial u_r}{\partial z} - \frac{u_\phi^2}{r}\right) = \\
-\frac{\partial p}{\partial r} + \gamma\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_r}{\partial c}\right) + \frac{1}{r^2}\frac{\partial^2 u_r}{\partial \phi^2} + \frac{\partial^2 u_r}{\partial z^2} - \frac{u_r}{r^2} - \frac{2}{r^2}\frac{\partial u_\phi}{\partial \phi}\right] + \rho g_r \\
\phi : \rho\left(\frac{\partial u_\phi}{\partial t} + u_r\frac{\partial u_\phi}{\partial r} + \frac{u_\phi}{r}\frac{\partial u_\phi}{\partial \phi} + u_z\frac{\partial u_\phi}{\partial z} + \frac{u_r u_\phi}{r}\right) = \\
z : \rho\left(\frac{\partial u_z}{\partial t} + u_r\frac{\partial u_z}{\partial r} + \frac{u_\phi}{r}\frac{\partial u_z}{\partial \phi} + u_z\frac{\partial u_z}{\partial z}\right) = \\
-\frac{\partial p}{\partial z} + \gamma\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_z}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u_z}{\partial \phi^2} + \frac{\partial^2 u_z}{\partial z^2}\right] + \rho g_z
\end{aligned} \tag{1}$$

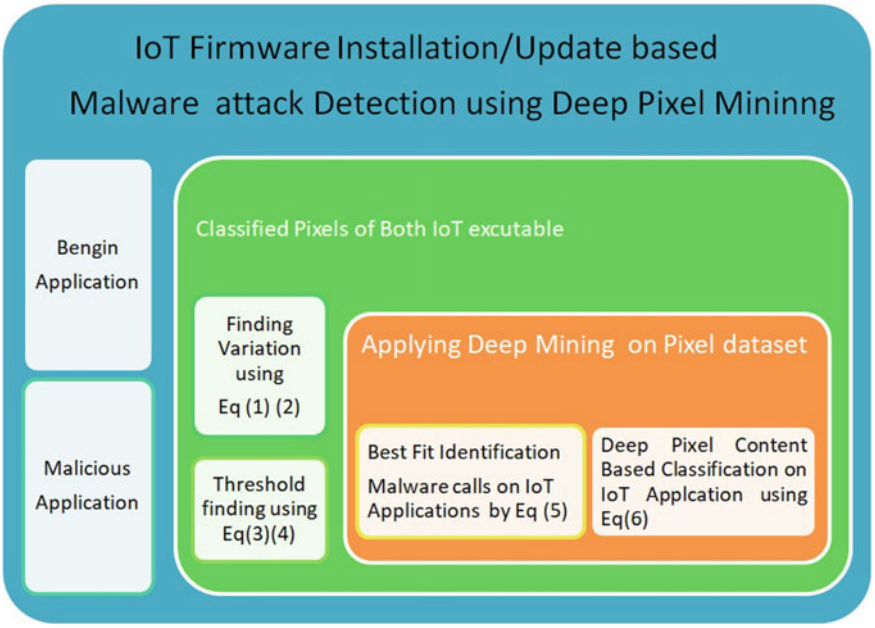


Fig. 1 Architecture of proposed higher order statistical parameters (FA-PCM)

The expectation parameter denoted by E and $E(x)$ is denoting mean value.

Equations (2–3) are used to calculate sample lopsidedness value from sample of N values.

$$\begin{gathered} \int_{-\infty}^{\infty} c^{-x^2} dx = \sqrt{\pi} \\ E(x) = a_0 + \sum_{n=1}^{\infty} \left(z_n \cos \frac{n\pi l}{R} + b_n \sin \frac{n\pi L}{R} \right) \\ x = \frac{-e \pm \sqrt{e^2 - 4dc}}{2d} \end{gathered} \tag{2}$$

where

$$\begin{aligned} \mathrm{E(c,\,d)} &= \sum_{i=0}^{j-1} {}^{k-1} \sum_{j=0}^{-j2\pi \left(c \frac{d}{H} + d \frac{c}{P} \right)} \\ \mathrm{e(c,\,d)} &= \frac{1}{CD} \sum_{i=0}^{j-1} \sum_{n=0}^{k-1} E(c,d) e^{k2\pi \left(c \frac{d}{P} + y \frac{c}{H} \right)} \end{aligned} \tag{3}$$

The weight of the tails is assessed by coefficient of variation in the dispersion. Throughout the distributions of high income generating, outliers and strong tail were noticed. In low correlation distributions, the E(c, d) coefficient is the flat top in the middle instead of the high point of the IoT executable pixel classification value. The calculated values are small because the proportions are symmetrical.

The γ_4 is denoting the random variable x of Leptokurtic in Eq. (4) [9, 11].

$$(x+y)G_\gamma(x,y) = 1 + \gamma \int_0^\infty dp \left(\frac{G_\gamma(p,y) - G_\gamma(x,y)}{p-x} + \frac{G_\gamma(x,y)}{1+p} \right) + \gamma \int_0^\infty dq \left(\frac{G_\gamma(x,q) - G_\lambda(x,y)}{q-y} + \frac{G_\gamma(x,y)}{1+q} \right) - \gamma^2 \int_0^\infty dp \int_0^\infty dq \frac{G_\gamma(x,y)G_\gamma(p,q) - G_\gamma(x,y)G_\gamma(p)}{(p-x)(q-y)} \tag{4}$$

An estimate of the Leptokurticis given in Eqs. (5–6)

$$\begin{aligned} P(a) &= \textstyle\sum_{i=0}^{i-1} f(c) e^{-k2\pi\left(c\frac{d}{K}\right)} \\ p(a) &= \tfrac{1}{K} \textstyle\sum_{k=0}^{C-1} P(c) e^{j2\pi\left(k\frac{d}{C}\right)} \end{aligned} \tag{5}$$

where

$$\begin{aligned} P(m) &= \sum_{m=0}^{K} x(2m) H_P^{2mn} + \sum_{m=0}^{P} x(2n+1) W_H^{(2a+1)n} \\ &= \sum_{m=0}^{2} x(2m) H_{\frac{P}{2}}^{mn} + P_H^b \sum_{m=0}^{\frac{P}{2}-1} x(2m+1) H_{\frac{P}{2}}^{mk} \end{aligned} \tag{6}$$

where the mean value is $\tilde{P}$, the fourth central moment is m_4, the third central moment is m_3, and the variance of the data is m_2.

By using FA-PCM technique, we found the malware regions present in the IoT executable. By making use of a method, the malware and normal regions were distinguished in the IoT grayscale image. The windows size of 60*60 was applied to two different images selected from dataset. The regions with malwares suspected are marked as malicious region. The randomly selected region without malware IoT file without malicious service calls and the randomly selected region with malicious service call of the benign IoT file was then calculated by histogram. From the malware and randomly chosen region, histogram is determined. The asymmetric region of histogram depicts malware. The symmetric region of histogram depicts malware.

In the regions were found asymmetric and heavy tail of distribution in suspected region, and it narrates lopsidedness and Leptokurtic values are high in the malware region.

That is

For each(x_1 x_2 …… x_n) (y_1 y_2 …… y_n)
{
IF yi with R+0
Then
Influence function f_B :
$f_{yB} = f_{B(x,y)}$

such that for y_i belongs to R+0
do
{ if x_i and y_i are close
//idendified with closeness(i.e., γ_4)
return1
else return 0
} while (all y_n defined) End while End if End For End For

With use of Eq. 2, the lopsidedness value that is third-order moment was calculated. With use of Eq. 5, the Leptokurtic value that is fourth-order moment was calculated. The lopsidedness and Leptokurtic values are taken, and the numerical analysis is done to find the normal and malware regions. FA-PCM was applied over the randomly selected six IoT malware images in the combination of three regions having malware and the three regions having normal. In Table 1 displays suspected regions of various lopsidedness and Leptokurtic values.

From Table 1, the FA-PCM analysis over IoT executable found malware regions having that lopsidedness value greater than 0.3 and Leptokurtic value greater than 5 are marked as suspected regions. From the analysis of various IoT executables using FA-PCM, the above threshold values are set. Hence, deep pixel mining classification was added to the FA-PCM's lopsidedness and Leptokurtic parameters which were used to define the malware in executables for IoT application.

Table 1 FA-PCM for dubious area

Region	lopsidedness ($\|I_k\|$)	Leptokurtic ($\|L_u\|$)
Dubious area-1	0.6729	6.312
Dubious area-2	0.7324	5.434
Dubious area-3	0.6133	4.112
Dubious area-4	0.3213	5.247
Dubious area-5	0.4102	5.731
Dubious area-6	0.5554	3.123

4 Results and Comparison

The Deep Binary Pattern Classification Mining (FA-PCM) based on band pass frequency method was applied over the grayscale images to identify the malware on the IoT executable. For analysis, 960 IoT executable files were taken, which were preprocessed and converted as the grayscale images [10, 12]. In the sample IoT executable dataset, 413 are malware and 547 are benign files containing 960. Table 3 displays the sample classification of different IoT malware families used for analysis [13, 14].

We used the malware detection methods suggested by Jiawei Su Han et al. [6] and Rakesh Singh Kunwar [2] to compare our proposed process.

Jiawei Su introduces a modern, compact approach for detecting malicious DDos in IoT systems. They retrieve malware pictures and use versatile algorithms to classify family groups, e.g., a grayscale image transformed from binary malware [13]. Such enhancements would enable the device proposed to be usable with even fewer IoT telecommunications machine resources [3]. The test data indicate that the proposed method will achieve 94.0% precision for the identification of goodware and DDoS malware, 81.8% specificity for the grouping of singular values and two major groups of malware [1].

Rakesh Singh Kunwar, malcontents, used the tools of steganalysis to insert the malicious payload with legal or innocent images [14]. Such harmful products are only a few lines of codes utilizing the weakness of apps. It provides the intruder with wireless monitoring to the unlawfully operating network. Within that background, the main objective is to recognize any application code present in the picture [4]. Then, this main part of the system is focused on the detection of the software and its results. In such sense, the classified data is seen in JPEG images which propagate via social media sites [7]. This approach gives the correct solution. Nevertheless, the two methods were not used to classify specific malware groups. We especially in comparison two techniques of experimental analysis with 960 IoT gray image scale files containing 313 malware picture files where its 397 spyware was detected using the developed model, 372 malware was identified by Jiawei Su entropy and 307 spyware was found by rakesh singh method. Jiawei Su and Rakesh Singh Kunwar have obtained a combination of 92, 121, and 191 high false processes from the proposed IoT malware attack detection on firmware installation/update using deep image pixel mining. Thus, the suggested approach has a true value ratio between 96.12% (397/413) and 0.09 (92/960) fps for each object, with the true value ratio of Jiawei Su being 90.07% (372/413) to 0.12 fps per image and the ratio of Rakesh Singh Kunwar being 0.19 Fps for the true value-ratio of 80.87% (334/413) for the photograph. Table 2 reviews the two existing forms of analysis with the new system.

Table 2 Comparison of true positive ratios of various detection methods with proposed FA-PCM

Community with malware	Analysis data taken	Suggested (FA-PCM)-ratio and TP ratio (%) identification of data	Jiawei System of Su and TP ratio (%) observed data	Rakesh Singh Kunwar found forms and TP ratio (%)
Mirai	64	59(92.18)	52(81.25)	43(67.18)
Silex	75	61(81.33)	57(76.00)	47(77.04)
BrickerBot	13	9(69.23)	7(53.84)	6(46.15)
NanoCore	57	50(87.71)	52(91.22)	42(73.68)
Pushdo	71	64(90.14)	54(76.05)	57(80.28)
zeus	41	36(87.80)	32(78.04)	35(85.36)
Carna botnet	31	26(83.87)	20(64.51)	19(61.29)
Viruses	61	53(86.88)	47(77.04)	40(65.57)

5 Conclusion

The malicious device is a big challenge to security for both IoT consumers and application driver's creators. This paper executes computer viruses in the context of IoT (as binary) and transforms them through gray-level image. A hacker tracker service method was built for effective malware identification in IoT executable files. Framework was used to categorize spyware files with frequency band pass. On the sub-image of the path, where descriptive and inferential statistics were measured, the rectangular patch of 60 × 60 marker was shifted. In the area hit by phishing and by studying normal and spyware files, the simulation parameters have been recommended as the sub-band of the significantly lower FA-PCM ideals. With 0.09% false positives image (FPI), the suggested approach obtained a TP ratio of 96.12%. The downside of the approach presented is the right identification of various kinds of malware. A few days later, the malware code searches vast amounts of data and requires time. The pause in opening a single file could in future be reduced without compromising the precision of detection.

References

1. Kancherla, K., & Mukkamala, S. (2013). Image visualization based malware detection. In *2013 IEEE Symposium on Computational Intelligence in Cyber Security (CICS)* (pp. 40–44). IEEE.
2. Kunwar, R. S., & Sharma, P. (2017). Framework to detect malicious codes embedded with JPEG images over social networking sites. In *International Conference on Innovations in information Embedded and Communication Systems (ICIIECS)* (pp. 1–4). IEEE. https://doi.org/10.1007/978-1-5090-3294-5.
3. Arul, E., & Manikandan, V. (2016). Malware detection using higher order statistical parameters. In S. Subramanian, R. Nadarajan, S. Rao, S. Sheen (Eds.), Digital Connectivity – Social

Impact. CSI 2016. *Communications in computer and information science* (Vol. 679). Singapore: Springer. https://doi.org/10.1007/978-981-10-3274-5_4.
4. Burguera, I., Zurutuza, U., & Nadjm-Tehrani, S. (2011). Crowdroid: Behavior-based malware detection system for android. In *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices* (pp. 15–26). ACM.
5. Shamili, A. S., Bauckhage, C., & Alpcan, T. (2010). Malware detection on mobile devices using distributed machine learning. In *2010 20th International Conference on Pattern Recognition (ICPR)* (pp. 4348–4351). IEEE.
6. Su, J., Vargas, D. V., Prasad, S. (2018). Lightweight classification of IoT malware based on image recognition. In *42nd IEEE International Conference on Computer Software & Applications* (pp. 664–669). https://doi.org/10.1109/compsac.2018.10315.
7. Wagner, M., Fischer, F., Luh, R., Haberson, A., Rind, A., Keim, D.A., Aigner, W., et al. (2015). A survey of visualization systems for malware analysis. In *EG Conference on Visualization (EuroVis)-STARs* (pp. 105–125).
8. Zhou, X., Pang, J., & Liang, G. (2017). Image classification for malware detection using extremely randomized trees. In *11th IEEE International Conference on Anti-counterfeiting, Security, and Identification (ASID)*. IEEE. https://doi.org/10.1109/icasid.2017.8285743.
9. Nataraj, L., Karthikeyan, S., Jacob, G., & Manjunath, B. S. (2011). Malware images: visualization and automatic classification. In *Proceedings of the 8th International Symposium on Visualization for Cyber Security,* VizSec '11. ACM (pp. 4:1–4:7). ISBN 978-1-4503-0679-9. New York. https://doi.org/10.1145/2016904.2016908.
10. Sajedi, H, & Jamzad, M. (2010). Selecting a reliable steganography method. In *2010 International Conference on Multimedia Computing and Information Technology (MCIT)* (pp. 69–72). IEEE.
11. Han, K. S., Lim, J. H., Kang, B., & Im, E. G. (2014). Malware analysis using visualized images and entropy graphs. *International Journal of Information Security, 14*(1). http://link.springer.com/article/10.1007%2Fs10207-014-0242-0.
12. Sun, H., Wang, X., Buyya, R., & Su, J. (2017). CloudEyes: Cloud based malware detection with reversible sketch for resource constrained internet of things (IoT) devices. *Software: Practice and Experience*, *47*(3), 421–441.
13. Alam, M. S., & Vuong, S. T. (2013). Random forest classification for detecting android malware. In *IEEE International Conference on Green Computing and Communications (GreenCom) and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing (iThings/CPSCom)* (pp. 663–669). IEEE.
14. Liao, S. -H., Chu, P. -H., & Hsiao, P. -Y. (2012). *Data mining techniques and applications.* Department of Management Sciences, Tamkang University. Elsevier Ltd. https://doi.org/10.1016/j.eswa.2012.02.063.
15. https://www.stratosphereips.org/datasets-iot.
16. https://ieee-dataport.org/open-access/iot-network-intrusion-dataset#files.
17. https://www.virussign.com/.

Life Saver Robotic Car for Accidental or Disaster Place Emergency Situation

Shanta Saha, Md. Omar Sharif Rajme, Bijan Paul, Mohammad Rifat Ahmmad Rashid, and Khan Raqib Mahmud

Abstract Our project aims to design a Rescue Car, which is an Arduino base manual control system that is controlled by smartphones. Nowadays, many incidents were happing in the world, which took many people's lives. For the lack of proper technology and proper equipment as well as the complex system, we cannot save people's life. We cannot ensure that accidents are not going to happen, but we can reduce the damage and save many lives. By developing this project, we try to prevent this thing; we are tried to help the people in a short time. This project is mainly focused on finding out the people in accidentals or disaster place. Not only finding the people but also gets an overview of that place to reduce the rescue time. In this paper, we prepare a system to rescue people in an emergency.

Keywords Arduino uno · Rescue car · Visualization · Smartphone app

1 Introduction

Various unwanted incidents are happening nowadays which are a threat to human life. The records say we can't stop this natural or human-created disaster, but we can reduce the number of losses of lives by using the proper technology. According to

S. Saha (✉) · Md. Omar Sharif Rajme · B. Paul · M. R. A. Rashid · K. R. Mahmud
University of Liberal Arts Bangladesh, Dhaka, Bangladesh
e-mail: shanta.saha.cse@ulab.edu.bd

Md. Omar Sharif Rajme
e-mail: omar.sharif.cse@ulab.edu.bd

B. Paul
e-mail: bijan.paul@ulab.edu.bd

M. R. A. Rashid
e-mail: rifat.ahmed@ulab.edu.bd

K. R. Mahmud
e-mail: raqib.mahmud@ulab.edu.bd

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_39

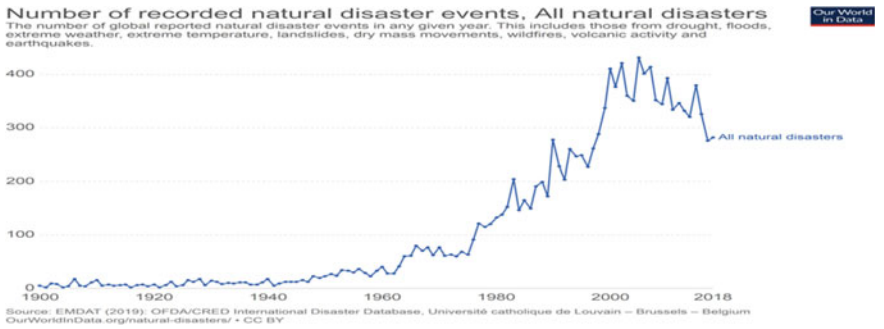


Fig. 1 World natural disaster [1]

Table 1 Bangladesh fire deaths record [2]

Country	Year	Deaths per 100,000 individuals
Bangladesh	2008	1.92296613049617
Bangladesh	2009	1.81293736505158
Bangladesh	2010	1.84862506546123
Bangladesh	2011	1.65546745422576
Bangladesh	2012	1.54293055814935
Bangladesh	2013	1.475276386832
Bangladesh	2014	1.39922492770893
Bangladesh	2015	1.33529433585401
Bangladesh	2016	1.2994316960941
Bangladesh	2017	1.27821563789698

the statistics we can say that the number of human death incident like a fire accident, earthquake is increasing day by day (Fig. 1 and Table 1).

In the last year 2019, many fire incidents happened in our country Bangladesh example of "Churihatta," "Banani" where many people lost their lives. If we prepare before it happened, then this huge type of incident did not happen. That's why we are trying to build a system that will help in those situations to rescue the people in a short time. This rescue car can easily be sent where people are not able to go. This car is controlled by our own design app which is specially designed for this purpose.

This paper is organized as follows in Sect. 2 describes related work, System design and project demo is demonstrated in Sects. 3, 4 and 5 describes User Satisfaction & Future concepts, Sect. 6 illustrate the Conclusion and at the end add the references.

2 Related Works

In the past, many people try to build an Arduino based rescue system. Nayantha KGD has implemented a Bluetooth Controlled Robot Car using Arduino. That can be controlled by using a mobile phone [3]. Selvam implemented a smartphone-based robotic control for surveillance applications that send live time video footage [4]. Lu et al. created a system Robot Control Design Based on smartphone [5]. A mobile robot navigation system based on a wireless control device has developed by Wek et al. [6]. In [6] Raja Lavanya has implemented an android application which uses Bluetooth for controlling Mobile Robot Temperature Monitoring System. Tanveer Ahsan developed an intelligent robots' cars using multiple motion control system and cloud based IoT services [7].

3 System Design

The design of the system has been kept as simple as possible. Prior to designing the project, a few issues like cost-effectiveness and simplicity of design have been kept in mind. Our aims design a system that can give these functionalities with a user-friendly interface:

(a) We develop an android application that uses our robot as a remote control and also fetches data from particular sensors to show the specific data. (b) We develop a robotic rescue car which can help to save people's life. (c) An android smartphone and the technology used in android devices can be used to interact with the embedded system. (d) The major component of this work consists of a Bluetooth module, Arduino UNO, Stepper motor driver, Ultrasonic sensor, PIR motion sensor, Temperature sensor. (e) Bluetooth module will connect with the microcontroller to interact with the user.

3.1 Project Components

See Table 2.

3.2 Components Description

Arduino UNO R3 It is an open-source microcontroller board based on the Microchip ATmega328P microcontroller. In this microcontroller board, interface is designed based on digital and analog input/output (I/O) pins. Using these pins, a broad range of expansion boards and other circuits are interfaced. Moreover, Arduino

Table 2 Description of project components

(1) A bunch of jumper wire	(2) TIP135 PNP transistor
(3) Arduino Uno R3	(4) L298N Stepper Motor Driver
(5) Bluetooth module HC-0	(6) Wheel
(7) Gear motor	(8) Servo motor SG90
(9) Sonar sensor (HC-SR04)	(10) HC-SR501 PIR motion sensor module
(11) Buzzer	(12) LM-35 temperature sensor
(13) 9 V battery	(14) Speaker
(15) Breadboard	

IDE (Integrated Development Environment) is presented to perform task which is connected via a type B USB cables [8, 9]. We used this microcontroller as an central processing unit for our robotic car.

L298N Stepper Motor Driver It is an integrated monolithic circuit in a 15-lead Multi-watt and PowerSO20 packages. It uses a high voltage, a high current full-bridge driver that designed to utilize standard TTL logic level sand drive inductive loads. For example, relays, solenoids, DC and stepping motors. We use this motor driver to control the rescue car direction.

Servo Motor It is a rotary actuator or linear actuator. It can provide precise control of angular or linear position, velocity and acceleration. We use this motor to control the angle of the sonar sensor.

Gear Motor It is used to converts direct current electrical power into mechanical power. We use this for the movement of our rescue car.

Bluetooth Module HC-05 The Bluetooth module is used as an interaction for data transfer and retrieval between systems. We are using the HC-05 Bluetooth module as a receiver and transmitter.

Sonar Sensor HC-SR04 This is used to measure the distance between itself and an object [10]. Following are the steps generated by HC-SR04:

- A high-frequency sound as a sonar signal is send by the transmitter (trig pin).
- When the signal finds out a specific object, then that signal reflected by that object.
- Finally. the reflected signal is received by The transmitter (echo pin).

In our project, it is used to measure the distance of an object and then send the data to smartphone using a microcontroller and Bluetooth to draw a map in an accidental place.

HC-SR501 PIR Motion Sensor Module The PIR motion sensor is a sensor that detects any motion in front of the sensor. We use this sensor to identify or looking for life by detecting motion.

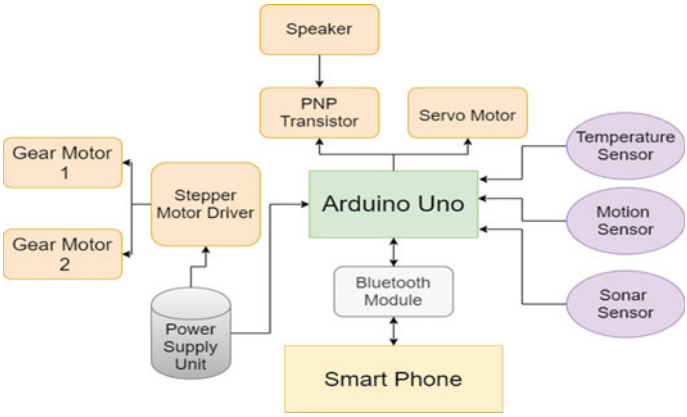


Fig. 2 Block diagram of life saver rescue car

LM-35 Temperature Sensor The LM35 temperature sensor is used to measure the temperature of any place. The measurement range of LM35 is between −55 and +150 °C. We use this temperature sensor to measure the temperature of a disaster place.

3.3 Block Diagram

In this section, we explored various sub-units of the system as illustrated in the block diagram of life Saver Rescue Car in Fig. 2. The system will consist of two main sections hardware and software. This block diagram is shown that the power supply unit provides the power in Arduino and the Stepper Motor Driver and other power of the component supply form the Arduino. All of the sensors are connected with the Arduino and then Arduino communicates with Smartphone with help of Bluetooth Module.

3.4 Software Section

Programming Language
The programming language used in Arduino which is a simplified version and compatible with both C and C++ programs. We write the whole code in the Arduino Integrated development environment (IDE) for smoothly running the rescue car as well as the other component. We use different types of Libraries such as “PCM.h” for speaker, “Servo.h” for servo motor, “SoftwareSerial.h” for Bluetooth module and many more. We kept the Arduino code as simple as possible and also bug-free.

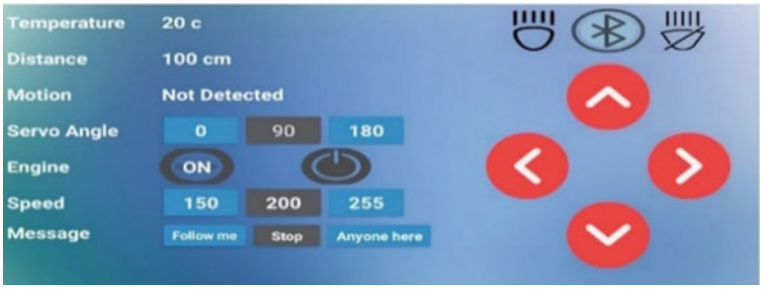


Fig. 3 Android application for controlling the rescue car

Android Application

In this project, we also develop our own Android application which is specially designed for controlling and collecting the sensor data from the Arduino. We kept this application as simple user-friendly as possible so that anyone can use this with a simple guide. By using this app, we can control the Robotic Rescue Car and measure the distance of any objects. We also turn on and off the headlight of our car so that we can continue our rescue mission at night. Furthermore, we can send some predefine audio that will help to communicate with the people. Most importantly our app collects all sensor data such as temperature, the distance of the object and also detects the motion. Suppose some blind people stuck in accidental place or the place is too dark to see anything that time motion sensor will help by detecting their motion. At the same time beeping the buzzer so that those people know someone is here to help or rescue them. We also added one important feature that is Visualization with the Smartphone camera by the screen mirroring technique. Here is a screenshot of our Android app (Fig. 3).

3.5 Working Principle

The working principle is kept as simple as possible. The working principle of our system has been described with the use of a circuit diagram (Fig. 4).

As seen from the circuit diagram a DC power supply is required to run the system. Whole system control by a microcontroller, where we are uploading our project preference program. The Bluetooth module transmits and receives the signal between microcontroller and smartphone. In the application, there have four directional arrow buttons. Those buttons are using to control the car movement. The upside arrow button sends a forward command to the microcontroller to move the car forward. The other 3 arrow buttons work the same way. In our application, there are also many buttons that are used for various purposes such as stop command button instantly stop the car, light on-off button controls the car headlight, and many more. There is also a temperature and sonar sensor which interacts with the microcontroller. After

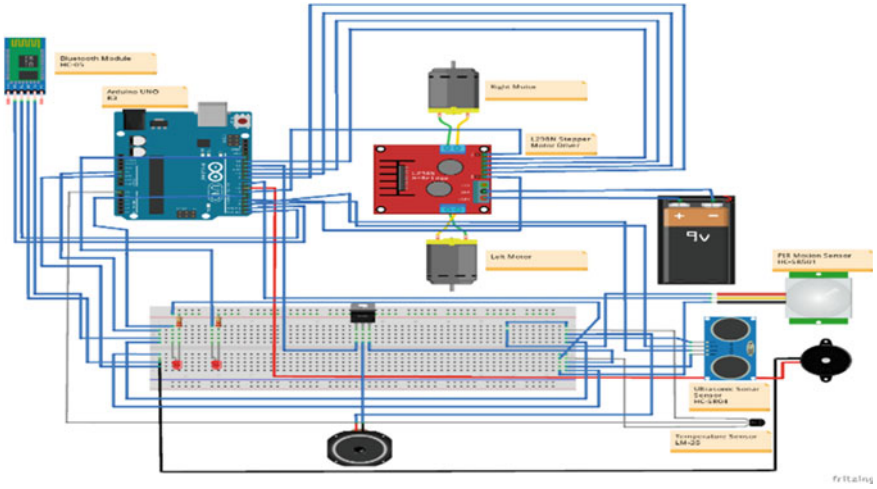


Fig. 4 Circuit diagram of the life saver robotic car

microcontroller collecting and processing the data, it simultaneously sends the data to the user smartphone. Buzzer and the motion sensor work together to find people. When motion (human) is detected then the buzzer will siren loudly. Using NPN transistor it transmits voice command to the speaker. The servo motor helps to move the sonar sensor 180°. From the application, we can control the sonar angle using a servo. We are also using a visualization system with the help of a third-party app "AirDroid." "AirDroid is a globally well-known application so there is no problem of trust. We don't even need an internet connection for the application to work. So there is no possibility of losing data to others. We need to use screen mirror technology between the two devices using a wireless access system. One device will set up in the Robot and another device will keep in user.

4 Project Demo

See Figs. 5, 6, 7 and 8.

5 User Satisfaction and Future Concept

After developed the system, we have collected user satisfaction doing survey around 30 people from a government organization in Bangladesh. The user satisfaction is shown Table 3.

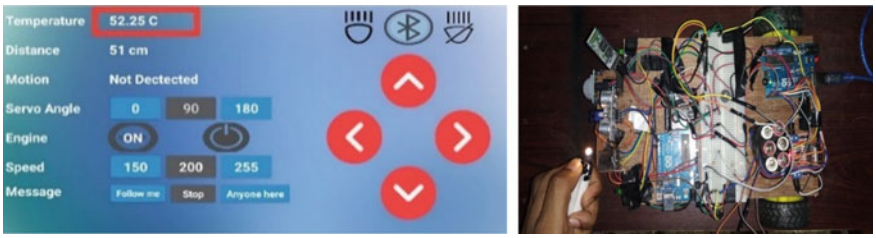


Fig. 5 Measuring temperature and sending the data to mobile application

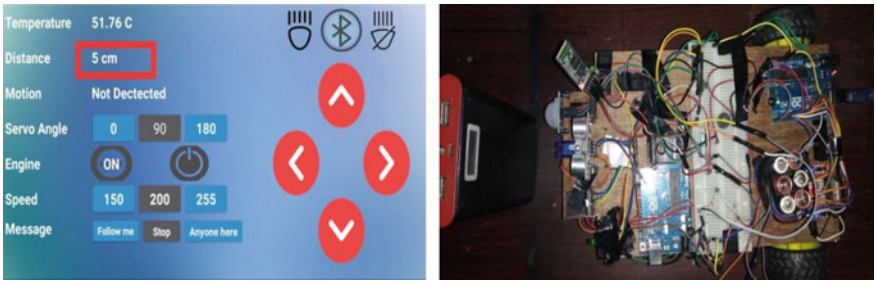


Fig. 6 Measuring the distance and sending the data to mobile application

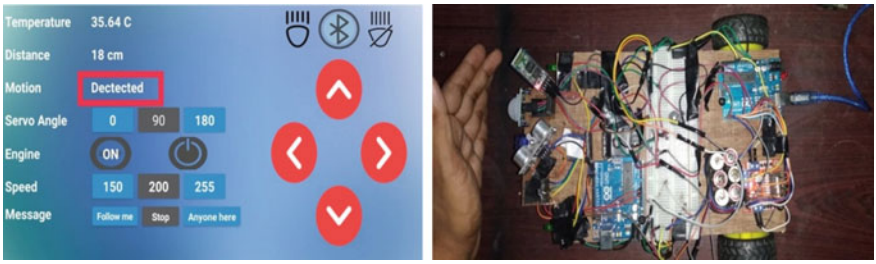


Fig. 7 Detected motion and sending the data to mobile application

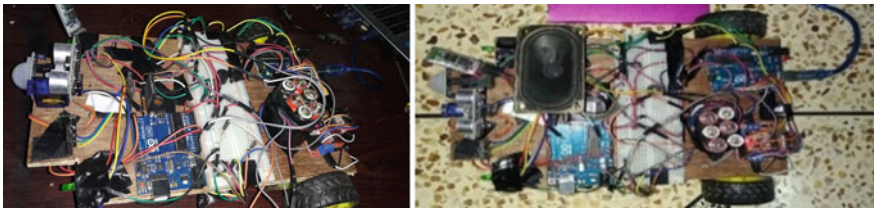


Fig. 8 Life saver rescue car with voice command

Table 3 User satisfaction level in percentage

Criteria	Satisfaction level (%)
Measuring temperature and sending the data	70
Measuring the distance and sending the data	65
Detected motion and sending the data	60
Life saver rescue	70
Voice command	60
Mobile app usability	80

Currently, our robot can rescue small items from a rescue place. In the future, we have a plan to improve this robot to gain advanced technology and system. For costing purposes, we are using screen mirror technology instead of a camera. We are using a Bluetooth module to interact between smartphone and robotic cars instead of wireless. We will also use in future GPS technology to point the specific location of a hostage. We have a plan for our invention to improve the Internet of Things (IoT) and improve security purposes.

6 Conclusion

Although our Robot is small in size it can be used as a rescue purpose with the help of some advanced technology and sensor. There have been many incidents in our country recently where many people have died due to lack of technology and proper equipment. So if we were ready before it happens then, we could save many lives as well as we can save out property before it damage. With our current technology and resources, we are not fully prepared for quick rescue management. So if we are properly prepared before anything happens we can save many lives as well as property through quick rescue management. Previously people try to build different type of rescue robot but in some certain point they are fail to fulfill people's proper demand. We try to solve this problem by adding different type of simple accessible technology at one place. So that our system can access by all type of people. This robot can be used to surveillance dangerous places. One of the main perspectives of our project is to reduce rescue time. At future by improving our project we will try to help much more people.

References

1. Brussels, U. C. D. L., *Number of Disasters (EMDAT(2019))* (International Disaster Database, Belgium, 2018).

2. G. B., & Network, D. C. (2017). *Deaths—Fire, heat, and hot substances*. United States: Institure for Health Metrics and Evaluation.
3. KGD, N. *Instrucatables*. [Online]. Available https://www.instructables.com/id/Bluetooth-Controlled-Robot-Car-Using-Arduino/. Accessed 03 05 2019.
4. Selvam, M. (2014). Smart phone based robotic control for surveillance applications. *International Journal of Research in Engineering and Technology (IJRET)*.
5. Lu, X., Liu, W., Wang, H., & Sun, Q. (2013). *Robot Control Design Based on Smartphone* (pp. 2820–2823). IEEE. 978-1-4673-1382.
6. Wek, P. S., Siew, Z. W., Wong, C. H., Chua, B. L., Teo, K. T. K. (2012). *Development of a Wireless Device Control Based Mobile Robot Navigation System.* IEEE.
7. Jenifer, T. M., Priyadharshini, T. S. V., Lavanya, R., & Pandian, S. R. (2013). Mobile robot temperature monitoring system controlled by Android application via Bluetooth. *International Journal on Advanced Computer Theory and Engineering (IJACTE)*.
8. Marques, C., Cristovao, J., Lima, P., Ribeiro, I., Ventura, R., & Frazao, J. (2006). RAPOSA: Semi-autonomous robot for rescue operations, intelligent robots and systems. In *IEEE/RSJ International Conference*.
9. Mishi, M. R., Bibi, R., & Ahsan, T. (2017). *Multiple Motion Control System of Robotic Car Based on IoT to Produce Cloud Service*. IEEE. Department of Computer Science and Engineering, International Islamic University Chittagong.
10. Tutorials, R. N. (2019). *Random nerd tutorials*. [Online]. Available https://randomnerdtutorials.com/complete-guide-for-ultrasonic-sensor-hc-sr04/. Accessed 03 05 2019.

Detection of Fat in Milk Using Photoconductivity and Color Detection Tecnique with Smart Billing

Vasudha V. Ayyannawar and Soumya R. Metri

Abstract Agriculture is a major part of India, and dairy bussiness is combined profit for the Indian business or the economy. Farmers supply their milk to the dairies and get the payment based on the purity of milk that they add per liters. As we know, nowadays, world has become more faster with trending requirements, and needs are fulfilled with more luxurious lifestyle adoption by the people; therefore, it is necessary to improve the present lifestyle of Indian farming and dairy management system. The various factors are like fat, classification of cow milk or buffalo milk, and generating the desired rate for the amount of fat calculated in the milk. So, the system calculates these parameters, and simultaneously the payment is done automatically. The system has two major part of module; one is kept at milk storage center, and the other is at billing counter. The microcontroller reads the data & sends it to the Android phone. Using Blynk application installed in the phone, the calculation of bill can be made for calculating daily payment. The system helps to detect the fat content in the milk and provide smart mobile application. The sensors are coherenced with the Arduino board and microcontroller. Cost of the system is also low to detect fatness in the milk sample. Using the Internet of Things (IOT) process, the industry can take the real-time readings of milk and rate to the government which helps to stop the illegal things happening nowadays and getting fair price to the farmers

Keywords Milk storage center (MSC) · Blynk · Arduino board · Microcontroller · IOT

1 Introduction

Depending upon the time and situation, there is a need to change the working system of the old modules like dairy farming or it may be agriculture. Firstly, it is required to calculate the amount of fat present in milk. Manual testing of fat and quantity is

V. V. Ayyannawar · S. R. Metri (✉)
Department of Computer Science and Engineering, Basaveshwara Engineering College (Autonomous), Bagalkot 587103, Karnataka, India
e-mail: soumya11.metri@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_40

time-consuming. Secondly, some dairies in villages do not have good milk testing equipments. In such condition, the milk sample can be tested, once the milk is free from fat which can take one to two hours. By that time, the milk packed in a plastic bags or bottles is unhealthy. Another reason is that since the process is done manually it can result to mistakes which is biggest loss to farmers. Therefore, to decrease the manual work and to get better result, there is a need to replace the existing system with a new system using the system where milk sample can be measured automatically and in low cost. As a result, huge no. of farmers are supplying their milk to the dairy. It is a responsibility of the dairies to get the quality of milk from each farmer and give fare amount to the farmers.

2 Related Work

To measure the fat content, the principle used is a optical scattering of light on the homogenized milk. To overcome the problems of present method of milk analyzing, here is the system that had spot calculation of fat present in milk. The fat is calculated by the light that is passed and scattered in the milk sample that is considered as a percent of fat in the milk. Sensor is a device that helps to classify the cow milk or buffalo milk and provides a corresponding output.

2.1 Fat Measurement

2.1.1 Gerber Method

In this method, the removal of fat is done by adding of sulfuric acid, and the fat present is seen directly on calibrated butyrometer. The fat is removed from the milk by centrifugal force, the reason to use sulfuric acid is to dissolve protein that comes around, and amyl alcohol is added to increase the fat separation.

Procedure

(1) Take 10 ml of acid in a pipette and add 10 ml of sulfuric acid into butyrometer. (2) Then, add 10.75 ml to the pipette and send the sample into butyrometer. (3) By using 1 ml pipette, add 1 ml of amyl alcohol and let it close and shake it well until no white bubbles or particles are seen. (4) Keep the butyrometer in a water for 5 to 10 min. (5) Now remove it and clean it with a dry napkin, and now place the two butyrometer opposite for 6 min. (6) Then, make it cool for 10 min. (7) Finally, the fat is left at a lower end of the pipette.

Fig. 1 Gerber method

2.1.2 CLR Method

This method states the law of floating solid object on a liquid when it is added in a milk. Therefore, in this method, the procedure is as follows: Add 70 ml of milk in a pipette and dip a lactometer into it, scale mark on the milk is displayed, and simultaneously, these readings are noted down in a farmer card or a book (Fig. 1).

3 Approach

Keeping in mind about the disadvantages of existing method of analyzing milk which is spoken previously, "detection of fat in milk using photoconductivity and color detection technique" is designed in such a way that its construction is easy and also familiar to operate.

Figure 2 shows block diagram of milk storage, and a list is shown below:

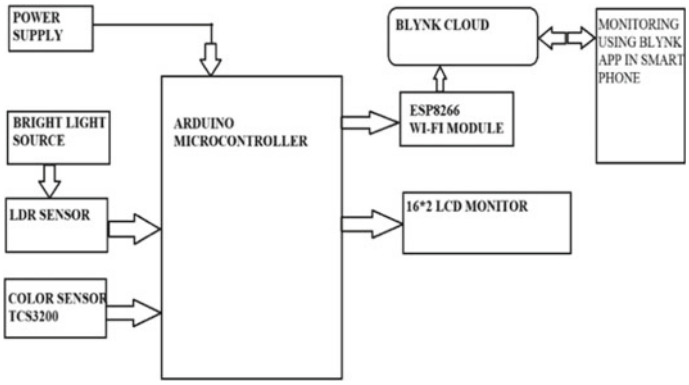


Fig. 2 Block diagram of entire module

- Aurdino microcontroller
- LCD display
- ESP-8266 Wi-Fi module
- Color sensor TCS3200
- LDR sensor
- Blynk cloud
  - Sensor block: The sensor used for detection of milk fat is LDR, and to distinguish between cow's milk and buffalo's milk, we use color sensor TCS3200.
  - Microcontroller: The role of microcontroller is to store the readings and by the mean time calculate the bill as well and display it on LCD screen. The microcontroller used here is ATMEGA 328.
  - LCD display: LCD is interfaced with microcontroller to show the readings of fat and rate per liter.
  - ESP-8266: This module is connected to Wi-Fi network to update the data in cloud.
  - Updated data is put on Blynk cloud and later controlled and monitored using a smart phone via a mobile application called 'Blynk'.

3.1 Working of Proposed System

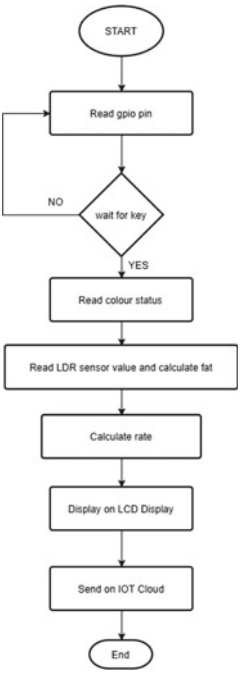
Milk tester is a method, which gives the result, that is, fat, based on the light scattered by the milk. A device called photoresistor whose resistance decreases when the incident light gets increases. It is a semiconductor material having high resistance.

It works on the principle of photoconductivity. When the light is fallen, the more number of electrons are released, which leads to increase in charge carrier those are holes. Thus, the results can be analyzed by the change caused in resistance, that is fat content in milk is shown.

For buffalo's milk, the fat present is 6–7%, and for cow milk it is 3–4%; if we do not get a fat in a range, the milk can be given with very low cost with no profit to the farmers.

Color sensor TCS3200 module is designed and used to detect whether the milk sample given under test is buffalo or cow's milk, by checking its color intensity, as we all know that cow's milk is slightly yellowish in color, whereas buffalo's milk is pure white. Based on this technique, we can easily identify whether it is cow's milk or buffalo's. Finally, the obtained data are displayed on LCD monitor and updated on Blynk cloud IOT platform, where data can be monitored through Internet itself.

3.2 Flowchart



3.3 Pin Configurations

ARDUINO UNO

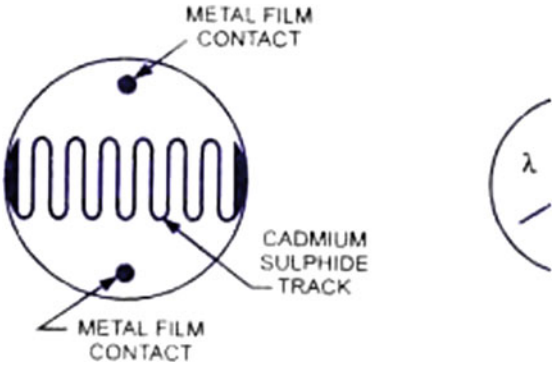
- MCU: Atmega 328
- Input voltage: 7–12 V
- Operating voltage: 5 V
- SRAM: 2 KB
- EPROM: 1 KB

3.4 Units

Arduino Uno board has two microcontrollers, ATMega 16 and ATMega 328. ATMega 16 has built-in RC phase oscillator with 2–8 MHz frequency, and it can generate its own oscillation. However, ATMega 328 is a open-source, hardware-abstracted, and inexpensive physical computing platform. It has built-in 32 K memory. The major components are USB connector, power port, microcontroller, analog input pin, digital pin, reset switch, crystal oscillator, and USB interface chip. It has 14 digital input–output pins in which six can be used as PWM outputs and six analog input pins.

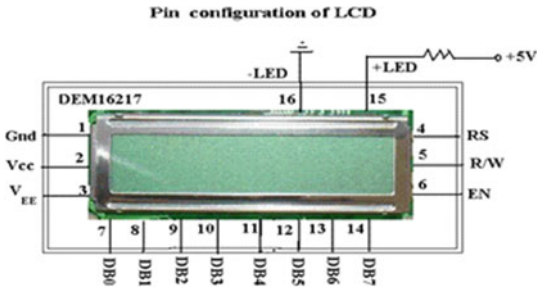
LDR Sensor

It is also known as photoresistor, it is a light-sensitive device, and they are made up of semiconductor having high resistance. When light falls that is when photons fall on device, the electrons in the valence bond of the semiconductor materials are excited to conduction band.



Working LDR is made up of most commonly cadmium sulfide. The zig-zag pattern shown helps in obtaining or getting accurate resistance. The zig-zag like area seperates the metal into two regions, the Ohmic contacts are done on both sides. Important to note is that this ohmic contact should be less as possible, so that resistance should get changed only based on the effect of light only. LDR needs only small power and voltage. Cadmium sulfide and binding materials are mixed, pressed, and sintered.

LCD Display(16*2)



Together millions of pixels has made a single display, combining of three sub-pixels that is red, blue, and green which form a single pixel. It is a mixture of two states of solid and liquid. LCD display are used in TVs, laptops, and cell phones. It is used to display the image. LCD is made with active display grid or it may be made up with passive display. The consumption of power is low that is the advantage of LCD.

Introduction to Blynk–IOT Plarform

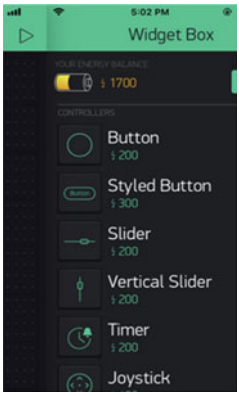
Blynk is designed for Internet of Things. Blynk can control or maintain the hardware device remotely. It can display sensor data, it can read data, it can store data, and it can visualize sensor data.

Blynk has the following three major components:

Blynk App: It helps to do good interfacing in our project with the help of widgets.

Blynk Server: It helps different communication between Android phone and the hardware. It is a open source which handles hundreds of devices. It provides secure cloud communication as well.

Blynk Library: With Blynk library, you can connect **over 400 hardware models** to the Blynk cloud. There are huge set of libraries which supports the blynk application.



4 Results and Analysis

See Figures 3, 4, 5, 6, 7, 8, 9, 10, and 11.

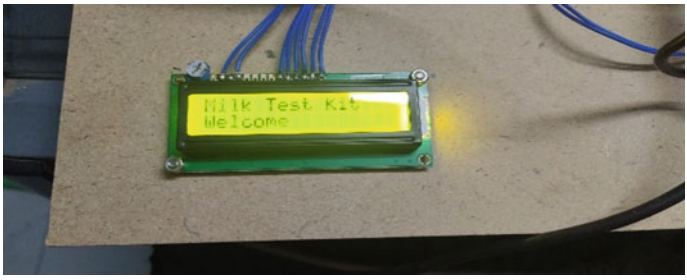


Fig. 3 LCD display screen

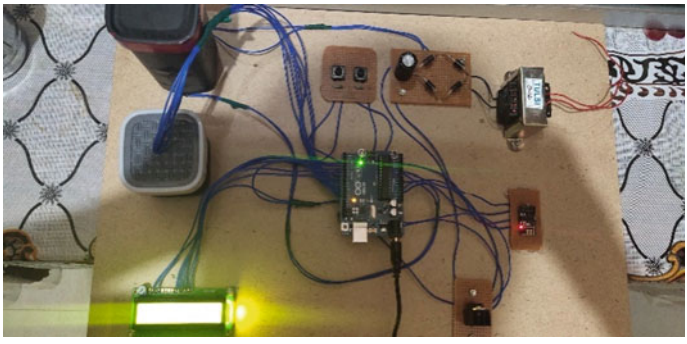


Fig. 4 Complete hardware

Fig. 5 Fat content in the milk

Fig. 6 Rate to be given for the milk

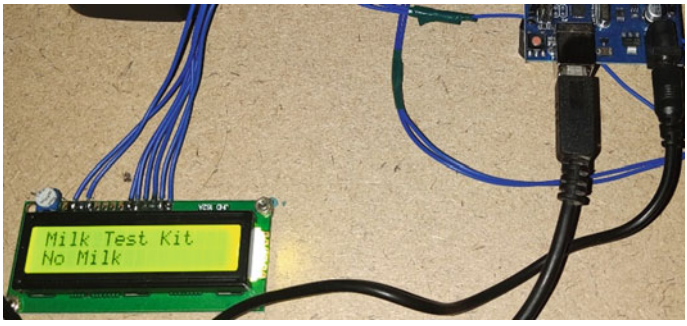


Fig. 7 When its completely water

5 Conclusion

In this paper, the system provides the ability to justify the quality of milk. The exact values of fat are displayed on LCD, simultaneously these values are sent over

Fig. 8 When its cow milk

Fig. 9 Fat content in the cow milk

the Internet, and anybody can retrieve the values over the Internet. Thus, with this accurate values, the farmers would get proper sale value of milk. The technology implanted in this system will definitely improve the system by giving fair price to farmers and also will minimize the corruption in the delivery system.

Fig. 10 Smart billing

Fig. 11 Rate and fat content display in a smart phone

References

1. Rupak Chakravarty, a paper on" IT at Milk Collection Centers in Cooperative Dairies: The National Dairy Development Board Experience", pp. 37–47.
2. Dabade, R.G. (2016). Milk analysing and billing system.*IJARIIE*,*2*(2). ISSN 2395-4396.
3. Draaijer, J. (2015).*Milk producer group resource book a practical guide to assist milk producer groups* (pp. 37–40).
4. Wolf, W.H. (1994).*Hardware-software co-design of embedded systems* (pp. 967–989). IEEE.
5. (2018). Milk quality control. In*International Center for Agriculture Reasearch in Dry areas (ICARDA).*

Feature Selection Optimization Using a Hybrid Genetic Algorithm

E. Padmalatha, S. Sailekhya, Saif Ali Athyaab, and J. Harsh Raj

Abstract The curse of dimensionality plays a vital role in data mining and pattern recognition applications. There are two methods which can address curse of dimensionality namely—feature reduction and feature selection (FS). The application of FS is such that it selects the most relevant subset of features with the less redundancy. Main objective of the proposed method is to manipulate irrelevant features and redundant features in (high–medium–low) dimensional data. We will aim to provide higher classification accuracy. In this proposed method, it is implemented with genetic algorithm. In this paper, we are working on different datasets to train and test our model.

Keywords Genetic algorithm · Feature selection · Optimization · Support vector machine · Linear regression

1 Introduction

Machine learning requires a lot of "good" data for obtaining accurate models. So, what is this so called "good" data? Simply put, it is data without excess noise which is cleaned and preprocessed as a suitable input to the ML model. In machine learning, the prediction model's accuracy always depends on noise in data. In the modern day, we have large infrastructure of the order of millions of IoT devices and sensors, due

E. Padmalatha (✉) · S. A. Athyaab · J. Harsh Raj
CBIT, Hyderabad, India
e-mail: epadmalatha_cse@cbit.ac.in

S. A. Athyaab
e-mail: athyaab@yahoo.co.in

J. Harsh Raj
e-mail: harshraj201999@gmai.com

S. Sailekhya
Prowess Enterprises, Vizag, India
e-mail: lekya.sheral@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_41

Table 1 Portuguese bank dataset description

Dataset characteristics	Multivariate	Number of instances	45211	Area	Business
Attribute characteristics	Real	Number of attributes	17	Date donated	2012-02-14
Associated tasks	Classification	Missing values	N/A	Number of Web hits	1,184,720

to which datasets are very rich in information. Because of this high-dimensional data, not always all the features may be required for prediction. By using feature engineering, only the relevant data can be applied to the learning models. Feature selection is a subset of feature engineering. It is commonly referred to as variable selection or attribute selection by ML researchers. Feature engineering reduces the training time, evaluation time and increases the accuracy [1]. In other words, the crux of the problem that we are trying to solve is "increase model accuracy and approach toward better performance [1, 2]."

2 Related Work

2.1 Introduction to Problem Domain Terminology

In data mining applications, feature selection (FS) is one of the fundamental data preprocessing techniques carried out before the actual analytics phase is even touched. FS is basically an optimization problem which may be solved combinatorically. For instance, a dataset with n features has a possible 2n solutions, therefore resulting in a tedious search space tree. Some popular search strategies include full (complete) search, greedy method, heuristic search, and random (stochastic) search. Nevertheless, these techniques lead to low accuracy, stagnating local optimum and high time complexity. As the old saying goes, "You need to know the problem inside out to truly do justice in solving it." That is what finally happens in FS. In majority of the cases, the technical and business teams need to sit out with a domain expert to solve the problem efficiently. Hence, FS may work on a better comprehension of the domain, by maintaining only the features with a differentiating factor and corresponding to some important measure. Describing noticeable patterns within the data to bring down dataset dimensionality is essential in the construction of a model to reduce the feature space as per our requirements.

Fig. 1 RMSE formula

$$\text{Fit}(g, C) = \text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

2.2 Existing Solutions

The existing solutions in feature selection can be applied for both supervised or unsupervised machine learning models. The first approach is "embedded methods." Here, the algorithms strive to achieve model fitting along with feature selection in parallel. The filter approach whereas tends to evaluate and finalize the important features purely based on the feature shape, i.e., characteristics. Alternatively, feature reduction is used in combination to feature selection or as a standalone. So what is the difference between the two? FS just chooses some features for model training from the entire dataset, i.e., excluding whole columns. FR transforms all the features into a lower dimension. Some FR methods remove features with missing values. Sklearn's feature selection module depends on a predefined threshold value—VarianceThreshold. It removes all features whose variance does not meet some threshold, eliminating all zero-variance features, i.e., features that have the same value in all samples. Another available FS [3] is bucketization or conversion of numeric columns into feature columns [4]. Let us consider 'age' as an example here. We can divide the data into discrete buckets, that is,

- Age 0–12 implies children
- Age 13–19 implies teenagers
- Age 20–50 implies workforce
- Age 50+ implies old/retired

3 Design of the Proposed Algorithm

3.1 Procedure for Developing a GA-SVM Model

The developed SVM model was trained with the selected kernel function after identifying training and testing data population. This was naturally done after performing a split on the main dataset (80% train, 20% test). The aforementioned SVM [5–7] was trained with each individual, and individual fitness function value$F$ was calculated. The metric used was the root-mean-square error (RMSE). The testing data RMSE was used to test the performance of the GA [8, 9]. In other words, the fitness function was actually the RMSE of the training data.

In this way, we can obtain a new categorical column from a numeric one. Although it is not necessary, it is recommended based on the current ML problem one is trying to solve (Fig. 1).

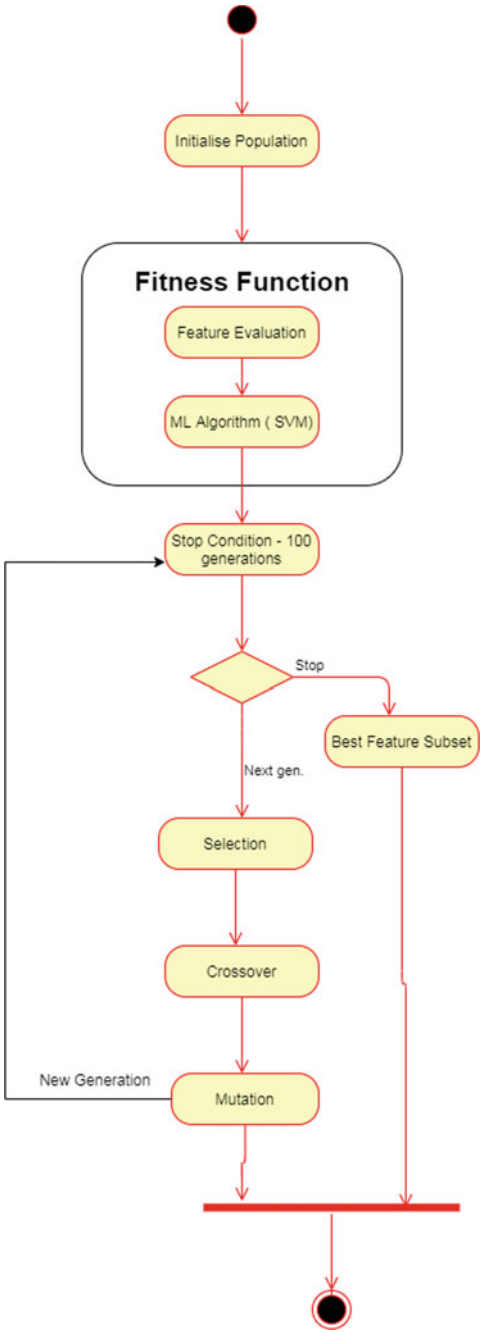


Fig. 2 Genetic algorithm activity diagram

Where y_i = real value of data set; $\hat{y}_i$ = the predictive value of the data set; and$n$ = proportion of the data settested (Fig. 2).

The algorithm goes like this. Upon obtaining the optimal individuals meeting the required fitness values or on reaching the maximum number of iterations, the individuals are returned (in this case, the best feature subset). Otherwise, the number of iterations is increased by 1. Selection, crossover, and mutation were applied to acquire upcoming populations, the SVM model was again trained, and the individual fitness function value was calculated. Then, conclusively, the optimized parameters were obtained. With the best matched parameters, the SVM model was tested with testing data, and the$y$ value was predicted.

3.2 Theoretical Foundation/Algorithms

Given below is the basic algorithm to be followed to aggregate SVM with GA, provided below are inputs and outputs requested (Fig. 3).

4 Implementation of the Proposed System

4.1 Flowchart of the Algorithm

See Fig. 4.

4.2 Dataset Description

The given data is connected with personal (one-to-one basis) marketing campaigns (phone calls) of a Portuguese bank. The end goal is to predict thar if the client will subscribe to a term deposit (variable$y$) with the bank after the marketing campaign (Table 1).

The marketing campaigns were entirely done through phone calls from dedicated call centers outsourced by the bank. Often, more than one contact to the same client was required (to convince the customer), in order to access if the product (bank term deposit) would turn out affirmative or not in terms of subscription. The only issue with such campaigns is the cost which the bank has to incur. So, the bank has to market the product targeting its customer individually (customer-segmentation).

Testing Process

What ML researchers fear the most is a model that simply responds with the labels of the samples that it has just seen. It would have a perfect score but ironically, it would

Fig. 3 Algorithm of proposed system

```
Input: S = (x1, x2. . . .. xn, y): Data set.
Where, x1.x2,..,xn are the features and y is the class
Psize: Population number.
Csize: Chromosome length.
Pc: Crossover probability.
Pm: Mutation probability.
T: Number of iterations.
Output: Best fitness and optimal feature subset OFS
1: Initialize algorithm parameters.
2: Load dataset S.
3: Sn =norm(S) // Normalizing the original data.
4: for i =1 to Psize do
5: for j =1 to Csize do
6: Pop(i, j) = round(rand) // Randomly initialize population.
7: end for(j)
8: end for(i)
9: for i =1 to Psize do
10: Subset = Sn(:, find(Pop(i, :) ==1))
11: end for(i)
12: t = 0 // initialize iteration number.
13: While maximum number of iterations is not meet T
14: for i =1 to Psize do
15: Fitvalue(i) = fit1(pop(i)) //calculate fitness of individual by SVM
16: sort(Fitvalue(i))
17: end for(i)
18: Pop = Elit(pop) //elitist preservation is applied
19: Chrom(t,Pop)=Chrom(t,Popc)/crossover operation
20: Chrom(t, Popc) = Chrom(t, Popm) //mutation operation
21: NewPop = Pop //produce new population
22: t = t + 1 23: end While(13)
24: OFS = S(:, Find(Bestindividual(1, :) ==1))
25: return Best fitness and optimal feature subset OFS
```

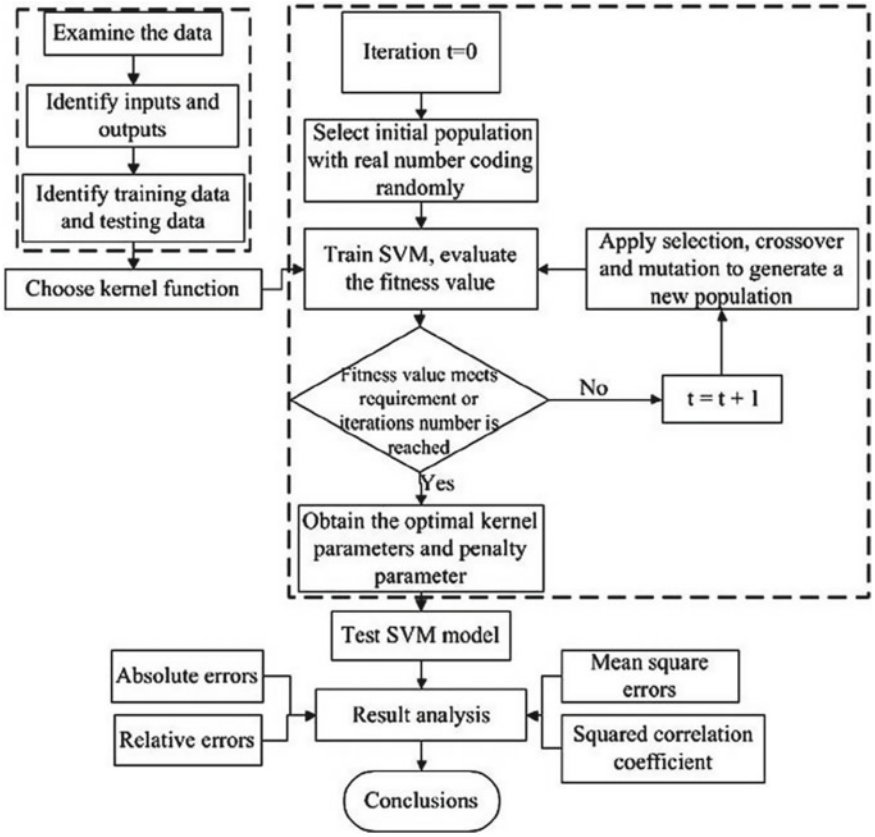


Fig. 4 Flowchart of GA with SVM

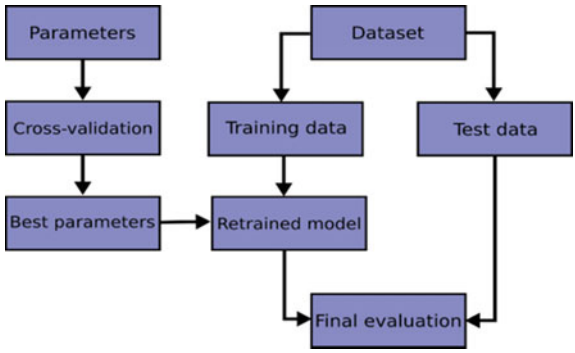


Fig. 5 Cross-validation workflow

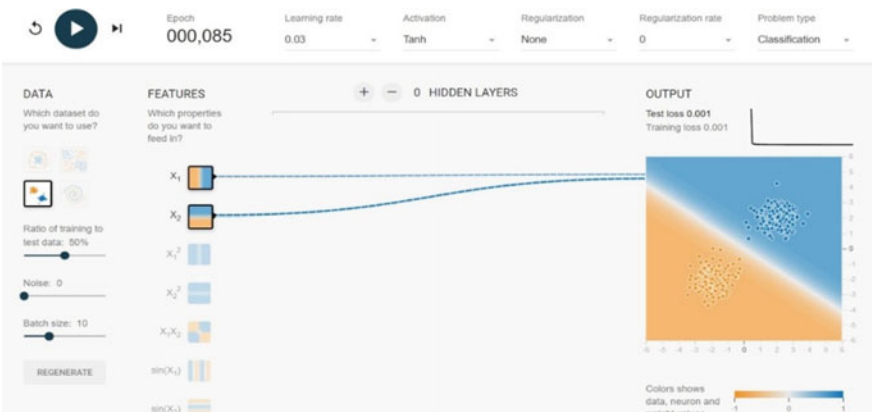


Fig. 6 TFP simulation using linear regression

fail to predict anything useful on yet-unseen data (just like a developing baby). This feared hindrance to a model's strength is termed as overfitting. A common practice is to keep aside a part of the dataset as a test set containing x_test and y_test. Given below is a flowchart of typical cross-validation workflow in model training. In our case, we use the 80-20 rule as mentioned before (Fig. 5).

5 Results/Outputs and Discussions

Initially using the linear regression model, we were able to get an 83+ accuracy level. We will use TensorFlow Playground, which is an intuitive Web tool to test out neural nets on datasets, to get a better understanding of what we are trying to accomplish here.

Let us consider this dataset which is probably closer to our bank dataset since our target output (y) is basically just a classification that is YES or NO. Using TFP's classification simulator and considering only two features, X_1 and X_2, we get a pretty good classifier with a training and test loss of only 0.001. This is an ideal case having negligible noise (Fig. 6).

Upon adding some noise and adding a few more features, we are now getting closer to an SVM model. We get a somewhat realistic view of our model's performance in Fig. 7. Notice that, the loss has increased to 0.012.

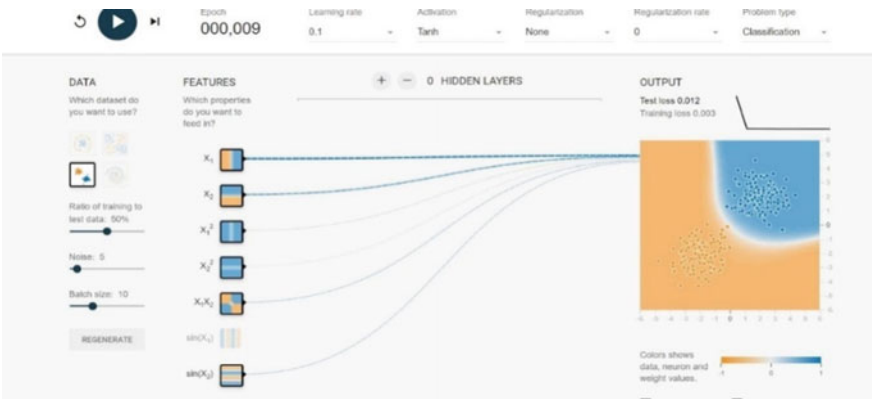


Fig. 7 TFP simulation with expected noise

6 Conclusions

The typical lifecycle followed by us during experimentation was as follows: first, evaluate the fitness value based on SVM; second, select a smaller feature subset generating or tending toward a higher accuracy to be included with the next generation; third, the remaining individuals are sent to crossover and mutation operations. Finally, the penultimate step is combining SVM and GA to aggregate them into FS for any dataset possible (large, small). After iterating through many generations, the algorithm reduces all the features to a "good features" subset. Upon discovering that the proposed method is capable of maintaining a pretty good balance on the feature subset size and classification accuracy, it may not be wrong to state that this approach can be applied in various research fields using ML classifiers such as computer vision, bioinformatics, and NLP. The future plan is set on embarking on a fresh search strategy and alternatives to ameliorate the skilfulness of the algorithm.

6.1 Limitations

Stochastic literally translates to "a natural tendency toward randomness." So, one has to keep that in mind before expecting high-order ambitious results. So, obviously amounting these algorithms to equality is a far-fetched fruitless effort. There is more than required amount of randomness which can lead to sub-optimum results.

Genetic algorithm is highly influenced by the initial population used. When expecting a wide diversity of feasible solutions, we must adhere to using datasets which have diverse feature values.

Since genetic algorithm belongs to a non-deterministic class of algorithms, the optimal solution, thus, may vary for each iteration of the running your algorithm for the very same input data. Closeness of the consequent solutions depends on the

convergence of the model which can be worked upon using hyper-parameter tuning concepts such as learning rate, epochs, and regularization. Lastly, we know that each dataset is different in its own way. Often, domain knowledge in the dataset being used is highly valued when working to improve the model accuracy, because only domain experts can tell which feature may be useful or which or not.

6.2 Recommendations and Future Work

GA being a very randomized algorithm was able to do a clean job combining with SVM. However, there is always scope for improvement. Delving deeper into fine tuning and getting closer to that 99% accuracy requires a deep understanding of deep neural networks and topics such as hyperparameter tuning, feature polynomials, hidden layers, activation function, and regularization. As stated above, it also requires domain experts, clean, and structured datasets to work with which may not be easily attainable in today's world. We aim to build further on this after acquiring the necessary skills and knowledge in the field.

References

1. Spolaor, N., Lorena, A. C., & Lee, H. D. (2011). Multi-objective genetic algorithm evaluation in feature selection. In*International Conference on Evolutionary MultiCriterion Optimization* (pp. 462–476).
2. Mao, Q., & Tsang, W. H. (2013). A feature selection method for multivariate performance measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence 35*(9), 2051–2063.
3. Fan, M., Hu, Q., & Zhu, W. (2014). Feature selection with test cost constraint.*International Journal of Approximate Reasoning*, *55*(1), 167–179.
4. Chakravarty, K., Das, D., Sinha, A., Konar, A. (2013). Feature selection by differential evolution algorithm—A case study in personnel identification.
5. Seera, M., & Lim, C. P. (2014). A hybrid intelligent system for medical data classification. *Expert Systems with Applications 41*(5), 2239–2249.
6. Chen, G., Chen, J. (2015).*A Novel Wrapper Method for Feature Selection and its Applications.* B. V.: Elsevier Science Publishers.
7. Goldberg. (1989).*Genetic Algorithms-in Search, Optimization and Machine Learning.* London: AddisonWesley Publishing Company Inc.
8. Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *Int. J. Prod. Econ., 133*(2), 662–676.
9. Dahea, W., & Fadewar, H. S. (2018). Multimodal biometric system: A review.*International Journal of Research in Advanced Engineering and Technology*,*4*(1), 25-31.
10. Xue, B., Zhang, M., Browne, W. N., & Yao, X. (2016). A survey on evolutionary computation approaches to feature selection. *IIEEE Transactions on Evolutionary Computation 20*(4), 606–626.
11. Kashef, S., & Nezamabadi-Pour, H. (2015). An advanced aco algorithm for feature subset selection. *Neurocomputing, 147,* 271–279.
12. Harvey, D. Y., & Todd, M. D. (2014). Automated feature design for numeric sequence classification by geneticprogramming. *IIEEE Transactions on Evolutionary Computation 19*(4), 474–489.

13. Bhadra, T., & Bandyopadhyay, S. (2015). Unsupervised feature selection using an improved version of differential evolution. *Expert Systems with Applications 42*(8), 4042–4053.
14. Brown, G., Pocock, A., Zhao, M. J., & Lujn, M. (2012). Conditional likelihood aximization: A unifying framework for information theoretic feature selection. *Journal of Machine Learning Research 13*(1), 27–66.
15. Dong, H., Li, T., Ding, R., & Sun, J. (2018). A novel hybrid genetic algorithm with granular information for feature selection and optimization.*Applied Soft Computing*,*65*, 33-46.
16. Mitra, P., Murthy, C. A., & Pal, S. K. (2002). Unsupervised feature selection using feature similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*(3), 301–312.
17. He, X., Cai, D., & Niyogi, P. (2005). Laplacian score for feature selection. In *International Conference on Neural Information Processing Systems* (pp. 507–514).
18. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.
19. Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: criteria of max-dependency, maxrelevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence 8,* 1226–1238.
20. Cai, D., Zhang, C., & He, X. (2010). Unsupervised feature selection for multi-cluster data. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 333–342).
21. Yu, L., & Liu, H. (2004). Efficient feature selection via analysis of relevance and redundancy. *Journal of Machine Learning Research 5*(12), 1205–1224.
22. Ruiz, R. (2006). Incremental wrapper-based gene selection from microarray data for cancer classification. *Pattern Recognition 39*(12), 2383–2392.

Medical Chatbot for Novel COVID-19

Fahad Mehfooz, Sakshi Jha, Sahil Singh, Shreya Saini, and Nidhi Sharma

Abstract Conversational agents or more universally known as the chatbot were industrialized to respond to user's queries in a particular domain. Chatbot would serve as a software delegate which enables a computer to converse with human via natural language. A chatbot is a human-like conversational character (Shaikh et al. Int J Eng Sci Comput 6:3117–3119, 2016 [1]. This technology was coined in 1960s, with the intention to impersonate a human (how he would reply to a particular situation) so that the user feels that he is talking to a real person and not a machine. Conversational agent that interacts with user's turn by turn using natural language (Shawar A, Atwell E (2005) ICAME J Int Comput Arch Mod Med English J 29, 5–24, 2005 [2]). The world of chatbot has seen much of the advance since the invention, and they have progressed from conventional rule-based chatbot to unorthodox AI-based chatbot. The chat agents are expert in their fields [3]. The prime focus of this paper is to show implementation of a retrieval-based chabot with voice support, and we will investigate other standing chatbot and how it is useful in helping the patients fetching all the necessary details about COVID-19.

Keywords Natural language processing · Retrieval · ELIZA · Simulation · Set of responses

F. Mehfooz (✉) · S. Jha · S. Singh · S. Saini · N. Sharma
GGSIPU, Delhi Technical Campus, Greater Noida, India
e-mail: fahad.mehfoooz@gmail.com

S. Jha
e-mail: sakshi.jha1019@gmail.com

S. Singh
e-mail: sahilsinghtheblackeyedpeas@gmail.com

S. Saini
e-mail: saini.shreya1998@gmail.com

N. Sharma
e-mail: nidhisharmait85@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_42

1 Introduction

The world has seen a drastic change in technology; the facilities today have completely altered. There was a time when people did not rely on technology for everything. Because of this change, there is a need for the use of conversational agents. Chabot changes the way of providing the services. Instead of using Web sites or installing another new applications, users could simply order a service via a chat interface [4]. Lately, the world has suffered from a calamity, which has caused vast wreckage, namely COVID-19. The fatality rate is saddening. Facts suggest that people are not really aware about the causes and symptoms of COVID-19. There already exist conversational agents that reply to people's queries on countless infections but no work or fewer work is done on chatbot for COVID-19; therefore, our enthusiasm behind this work is to create a chabot that would answer all the basic enquiries and the quandaries faced by anybody who seeks help from this chatbot and give as much as possible to make people avail all the necessary help they need. It would answer queries like "Is traveling safe?", " what could be the possible symptoms of COVID-19," etc. Also, mostly chatbot is text-based, and this chatbot completely works on voice support.

2 Existing Systems

Chatbots can be categorized on the basis of their input processing and method of output generation as:

2.1 Rule-Based

Rule-based chatbots are trained on a given set of defined rules to generate suitable responses for user's queries. When the number of possible outcomes is fixed and scenarios are imaginable in number [5], such chatbots are very useful. Rule-based chatbots prevent nonsense responses, grammatical errors and spelling mistakes which can reduce user experience. If the chatbot is made with help of an expert of the subject on which the chatbot is created, it will be more accurate and reliable. As under the expert's guidance, a chatbot can reach its goal better and efficiently. For businesses that deal with straight questions that do not require much intelligence and customized answers [6], such chatbots are usually preferred. Two popular examples of chatbots using rule-based approach are:

ELIZA. Joseph Weizenbaum created ELIZA, which gives a computer the ability to have a natural language conversation with a human being. ELIZA program was made to simulate a psychiatrist's discussion with patients [7]. When the user is conversing with ELIZA, the user will type a normal statement. ELIZA will then

give a typed response according to the statement written by the user after analyzing it. Control is then again with the user [8]. Weizenbaum's famous simulation of a psychotherapist [9] gave inspiration for the creation of many chatbots in the world.

Parry. Psychiatrist Kenneth Colby created PARRY, an example of early chatbot created after ELIZA was made. PARRY was developed to simulate a paranoid patient [10]. Colby considered PARRY as a tool for studying the paranoia's nature and ELIZA as a clinical agent that can hourly manage hundreds of patients. It has also been claimed that PARRY, a simulation of paranoid thinking [11], had passed something that could rightfully be called a legitimate Turing test [7] that ELIZA could not.

Also, our proposed system uses this approach for greetings. The rule-based approach has its benefits, but it is also limited by the set of rules. It lacks intelligence and users might have questions that it is not trained for. Due to this reason and to make chatbots human-like, people started using machine learning to make chatbots.

2.2 Self-learning-Based

Chatbots following this approach use machine learning and artificial intelligence technologies for learning from their previous conversations for chatting. Self-learning chatbots are useful when users have different types of queries that are not restricted by rules. Chatbots following self-learning approach can be further categorized as:

Retrieval-Based. A retrieval-based chatbot ranks data and selects responses instead of a new one. The output will be the best ranked from the set of responses that are already defined and finite. The chatbot is trained to do so and is successful most of the time. Retrieval-based bots are the most common type of chatbots that are created. Such an approach is probably the best option for bots that have a goal. Our proposed system majorly uses this type of self-learning for the goal of making people aware of COVID-19 (Fig. 1).

Generation-Based. A generation-based chatbot generates responses instead of selecting them from the underlying model [12]. Chatbot using this model is trained with a vast set of early conversations instead of predefined responses which may not always be available. Such a chatbot has a response for any query, but it may not always make sense and might have grammatical and spelling errors due to which it needs more research (Fig. 2).

With a detailed comparison between retrieval-based and generation-based as shown below, we can see why our proposed system is retrieval-based.

3 Generation-Based Versus Retrieval-Based

Generative bots are used widely, but retrieval-based bots provide more enhanced speech processing and extracting measures, thereby helping in generation of almost accurate results and learned responses. They learn from the set of responses stored

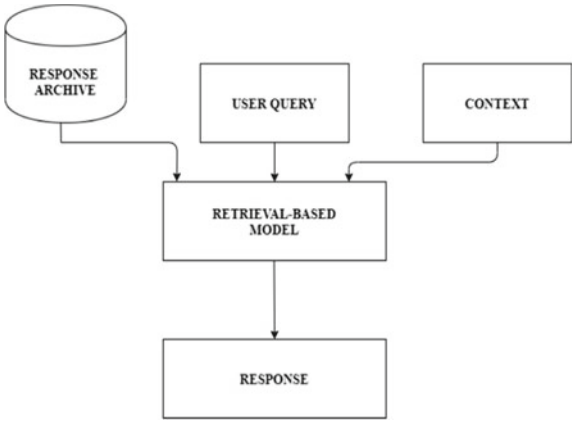


Fig. 1 Flow of retrieval-based chatbot

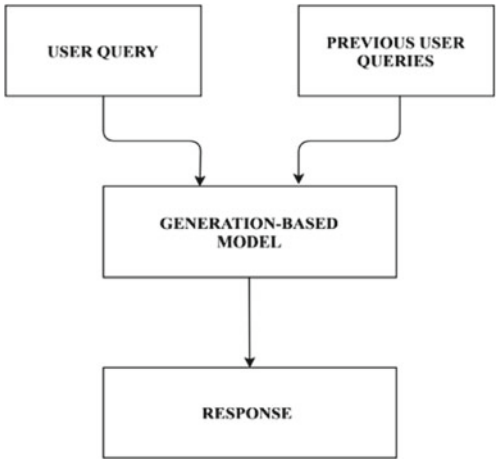


Fig. 2 Flow of generation-based chatbot

and try to develop a learning behavior based on the previous stored quoted responses. This increases the fluency, accuracy and diversified nature of the output. Response suggestion is an important for building computer–human-aided conversation systems [13].

Neural architecture had been used in the previous studies which improvised the nature of responses generated. Multi-contexture encoding was pulled over single text encoding enhancing the performance up to significant level. The majority deployed chatbots are complete retrieval model and have some limitations for the facts too. These bots already have proper answer and require no manipulation, framed with almost no grammatical errors. But, in some cases, these bots fail to generate those responses even properly framed question by the user. Retrieval is based on database stored in the machine, and if the database is efficiently arranged using suitable and emphasized algorithms, then some queries can be solved more accurately.

They are incapable to refer to contextual entity datasets like names mentioned previously in the conversation. While, secondly, they lean on programmed framed repository of predefined responses [14], they do not exhibit grammatical mistakes, because they pick from the user responses and offer a satisfied user experience who interacts with these bots.

While, in the generative-based bots, the answer and the response are struck between growing and learning states which means that there is more work yet to be done to make these family of bots more better reliable entity and generate almost user-friendly, self-learnt responses. In generative bots, the conversation looks just like real-life conversation between humans. Generative bots need to be upgraded every time with new AI methods to generate a powerful tool in the respective field usage.

The limitation set of generative type involves training and testing of large data sets which makes it cumbersome approach; that is why, mainly, all the bots are based on the retrieval-based approach. Currently, the main application of these bots is solely in field of e-commerce, education and real-time mobile-based applications.

In modern growing world of today, various sites and methods are available that provide users with no prior knowledge to build and develop their very own friendly bots. Those bots would be comprised of online developer modules which can be interim connected via Internet services by defining handful of parameters; that is why, retrieval-based bots are considerate in these scenarios. In our developed application on "COVID-19", the collective efficient dataset is more viable to retrieval process because it mainly focuses on highlighting the quick response and providing the necessary information to the user even unknown of its deadly affects, whereas generative approach will not be as effective in this case.

But using the learning and the principles of deep learning, both methods can be modified for purpose-based requirement in the future.

4 Proposed System

In this paper, we highlight the speech-based as well as a text-based medical chatbot. The recent outbreak of COVID-19 brings light to the fact that a majority of the populous in oblivious of the information required to deal with it. The chatbot is designed to address the frequently answered queries like: "How can I protect myself from coronavirus?", "What is coronavirus?", "Symptoms of coronavirus" and many more. So, this chatbot will not only solve these queries but also give you some important news regarding the coronavirus. Not only will it make people aware of the basic information about COVID-19 but also help to reduce the chances of being infected. This system is based on natural language processing, speech recognition systems and machine learning (Fig. 3).

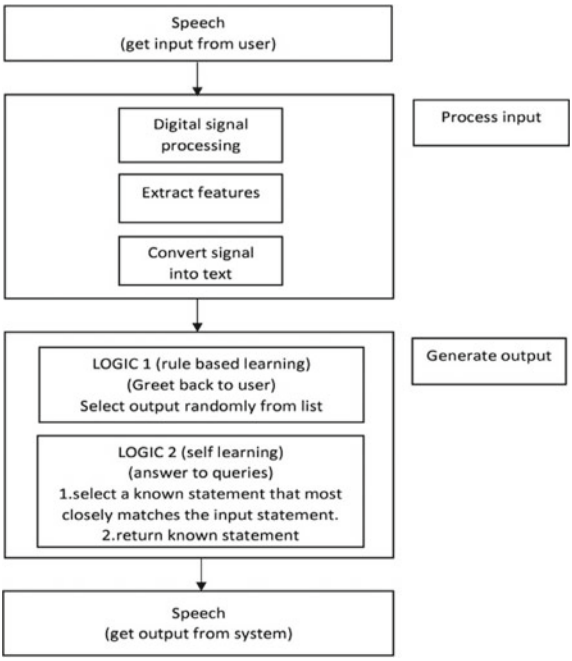


Fig. 3 Architecture of medical chatbot

5 Implementation

This chatbot is based on two types of learning algorithms: self-learning and rule-based learning. This system takes input in the form of speech using a microphone from the user. After taking a voice signal as an input, the system starts processing the voice signal and converts it into text data. Speech analysis can be broken down into given stages:

(i) Voice signal recognition and convert the signal into text
(ii) Text processing
(iii) Convert text into voice signal as a response.

Firstly, the user passes their query through a microphone. This system extracts features from the input signal. Specific instructions can be used to read the input signal and then to convert it into text. [15] Now, the given text is handled using natural language processing. In this, it carries out different natural language processing methods to sort out the required text by using various techniques like shallow parsing, deep parsing, semantic role labeling, conference relations and named-entity recognition. [16] Deep parsing technique is a technique in which the chatbot takes an input statement from the user and performs tokenization and segmentation [6]. The text input is split using word tokenizer to extract the keyword and eliminate stop word. Using cosine similarity() method, system measures similarity score and tries to find an answer from the dataset which is more similar to the user's query. Finally, chatbot

can be able to give an accurate, short and concise quick response to the user. The conversation between the user and the system can be done by both typing text and using voice. In the last step, it converts response text into a verbal response. Now, the text is analyzed before passing it to the linguistic model and converted into spoken words.

6 Conclusion

COVID-19 is a deadly virus which has taken world by storm. COVID-19 devastated USA, Spain, Italy, Germany and various other nations across the globe originating from Wuhan, China. Most people start to feel frequent fever, coughing, illness and sneezing. It is more fatal for old people and to those who have diabetes, asthma or any other kind of respiratory syndrome.

People are unaware of various measures of its prevention. Its cure is still in search, a kind of global health emergency, and awareness campaign is born like never before.

This project focuses on this serious concern, using knowledge on the bots created before, and it is just an attempt to provide a simple and yet powerful voice-based bot. It is based on the concept of machine learning combined with the real-world entity to highlight and aware people about the prevention and seriousness of this deadly virus.

References

1. Shaikh, A., Phalke, G., Patil, P., Bhosale, S., & Raghatwan, J. (2016). A survey on chatbot conversational systems. *International Journal of Engineering Science and Computing, 6,* 3117–3119.
2. Shawar, B, A., & Atwell, E. (2005). A chatbot system as a tool to animate a corpus. *ICAME Journal: International Computer Archive of Modern and Medieval English Journal, 29*, 5–24. ISSN 0801-5775.
3. Behera, B. (n.d). *Chappie—A semi-automatic Intelligent Chatbot* (pp. 1–5). https://www.cse.iitb.ac.in/~bibek/WriteUP_2016.pdf.
4. Kříž, J. B. (2017). *Chatbot for laundry and dry cleaning service* (pp. 1–61). https://is.muni.cz/th/eggbb/chatbot-diploma-thesis.pdf.
5. Nimavat, K., & Champaneria, T. (2017). Chatbots: An overview types, architecture, tools and future possibilities. *International Journal for Scientific Research & Development, 5,* 1019–1026.
6. Dua, A. (2019, May 29). Rule-based bots or Ai bots? Which one do you prefer? Retrieved from https://dataconomy.com/2019/05/rule-based-bots-or-ai-bots-which-one-do-you-prefer/.
7. French, R. (2000). The turing test: The first 50 years. *Cognitive Sciences, 4*, 115–122.
8. Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9,* 36–45.
9. Pruijt, H. (2011). *Social interaction with computers an interpretation of Weizenbaum's ELIZA and Her Heritage* (Vol. 24, pp. 516–523). SAGE Publications.
10. AbuShawar, B., & Atwell, E. (2015). ALICE Chatbot: Trials and outputs. *Computación y Sistemas, 19,* 625–632.

11. Colby, K. (1999). Human-computer conversation in a cognitive therapy program. In Y. Wilks (Ed.) *Machine conversations* (pp. 9–19). Kluwer. https://doi.org/10.1007/978-1-4757-56876_3.
12. Almansor, E. H., & Hussain, F. K. (2020). *Survey on intelligent Chatbots: State-of-the-art and future research directions* (pp. 534–543).
13. Swanson, K., Yu, L., Fox, C., Wohlwend, J., & Lei, T. (2019). Building a production model for retrieval-based Chatbots. *Computational Linguistics*, 32–41.
14. Veglis, A &Maniou, T, A. (2019). *Chatbots on the rise: A new narrative in journalism* (Vol. 7, pp. 1–6). Redfame Publishing.
15. Abdul-Kader, S. A., & Woods, J. (2015). Survey on Chatbot design techniques in speech conversation systems. *International Journal of ADVANCED Computer Science and Applications, 6,* 72–80.
16. Nair, G., Johnson, S., Verma, A., & Aarthi, S. (2019). A novel approach to understand the future of Chatbots for Medical assistance. *International Journal of Engineering and Advanced Technology (IJEAT), 8,* 1557–1562.

Design of Lower Limb Exoskeleton

Abhishek A. Nimje, Atharva P. Patil, and Dipti Y. Sakhare

Abstract This exoskeleton puts forth a unique design of robotic wearable devices that is ready to provide suitable gait rehabilitation to the elder ones or people after sustaining injury or disability. This system helps physical therapists to guide their patients with gait training programs being even far away from them using the prototype. This exoskeleton consists of a wearable device on either leg, which is controlled by a microcontroller, and the movement of the exoskeleton is made possible by a linear actuator which helps the user (patient) to lift and move the limb in the swing phase of walking. The actuator will simultaneously make the user walk within its domain. While operating the system, the user has to keep adding force to activate the force sensor. On the other hand, this wearable will also be beneficial for those people who find it difficult to walk in inclined planes. The exoskeleton has hip and knee movements powered in a single plane.

Keywords Wearable devices · Gait · Rehabilitation · Linear actuator · Force sensor

1 Introduction

According to a recent survey conducted by Census India says that 2.1% of our total population is physically challenged and of that 0.6%, i.e., 6,105,477 of people, are not able to walk properly. In a developed country like the USA, out of 50 paralyzed people, only one is being recovered and the rest are disabled for the rest of their life.

A. A. Nimje (✉) · A. P. Patil · D. Y. Sakhare
MIT Academy of Engineering, Dehu Phata, Alandi (D), Maharashtra 412105, India
e-mail: animje@mitaoe.ac.in

A. P. Patil
e-mail: appatil@mitaoe.ac.in

D. Y. Sakhare
e-mail: dysakhare@etx.maepune.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_43

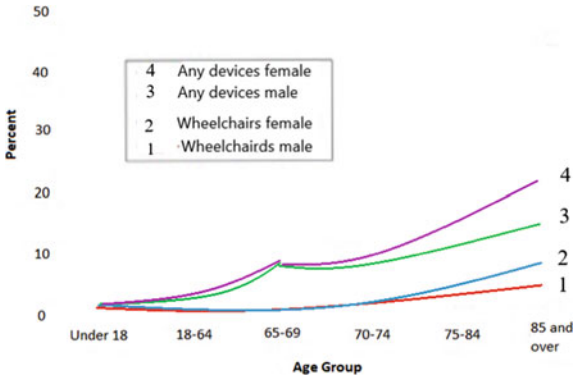


Fig. 1 Age group versus percentage wheelchair users

Henceforth, we developed a wearable device for those people who are disabled or paralyzed for a lifetime.

The methods include passive outhouses, exoskeletons, and physiotherapies. Exoskeleton comprises a light wearable brace support system, which integrates linear actuators at the hip and knee joints, rechargeable batteries, a couple of sensors, and a microcontroller. Every age group has some number of individuals who take the help of wheelchairs for their daily activity as shown in Fig. 1. Therefore, we have designed our exoskeleton in a way that it can be used for a wide range of people, and the width and height of the exoskeleton can be adjustable according to individual users. To walk smoothly and comfortably, these are the factors that play an important role: sufficient torque, sufficient speed, linear processing, parameter tuning, response time, and enough force to make a sensor to process signals to the microcontroller.

2 Literature Survey

The first research paper we referred talks about Cyberdyne's Hybrid Assistive Leg (HAL) developed by Japan's Tsukuda University. This exoskeleton has various versions of HAL 3 which mainly focuses on the lower limb and movement of the user, and another one is HAL 5 who helps the patients to move their whole body which is the arms, legs, and torso. The brain sends a signal to move the parts such as legs; then, sensors sense these signals sent to a robotic autonomous control system named cybernic controller which stimulates and controls the speed of exoskeleton [1].

This research paper talks about the simulation of exoskeleton and walk balance during both swing and stance phases. The simulation is done on the 3D model of the exoskeleton by using MATLAB/SimMechanics. The simulation result shows that moving forward is only possible by providing angular momentum to the hip joint and well as the knee joint [2].

In this section, they have discussed a commercially available exoskeleton named ReWalk which costs around $85,000 developed by ReWalk Robotics Ltd. This company has developed its two versions: ReWalk I which is usable for study and physiotherapy of the patient and the other one is ReWalk P which helps the patients for their daily activity. It is powered by a battery and operated by using a tilt sensor, and whenever the patient bends a little for walking, the sensor gets activated. The only drawback of this exoskeleton is its weighs that is 23.3 kg and it is difficult for every patient to afford this exoskeleton as its expensive [3].

Another exoskeleton was developed by Exo-Bionics, and this exoskeleton is pneumatically powered; therefore, it weighs around 20 kg. It has force sensors and motion sensors that help the exoskeleton to take the input and process the signals. This exoskeleton has a maximum speed of 3.2 km/h and has 6 h of battery life when used at its full capacity. The limitations to this exoskeleton are that users whose weight up to 90 kg and whose height is between 5 ft 2 in. and 6 ft 4 in. can only use the system [4].

3 Simulation Process

The exoskeleton can be worn in any of the legs just by changing the shoe. If the system was earlier installed and worn on the left leg, now the user just has to change the shoe and wear it on another foot. The system is intended and designed for persons with lower limb disabilities who have suffered injuries in the spinal cord. The system cannot keep balance and control, so the user should always be supported by the crutches (balancing equipment) (Fig. 2).

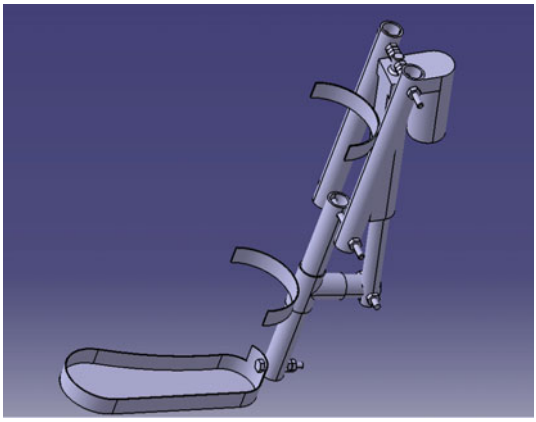


Fig. 2 3D model of exoskeleton including linear actuator

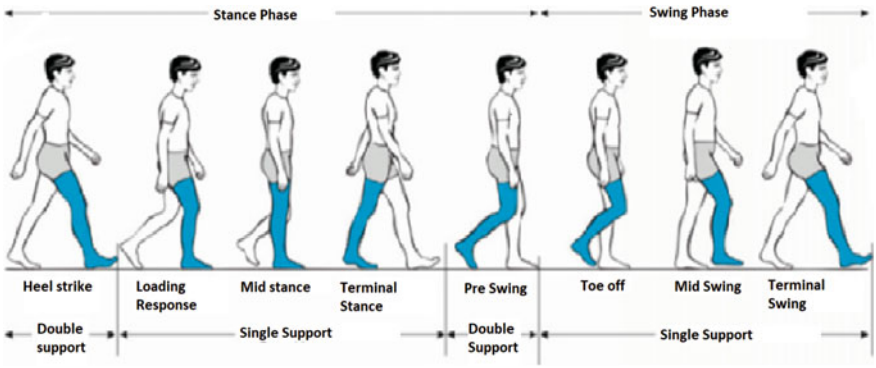


Fig. 3 Stance and swing phases [5]

3.1 Walking Simulation

The gait of a person is possible only if they follow rules of gait which include stance and swing phase.

Stance phase: This phase consists of the time when the foot is on the bottom and accounts for 60% of the gait cycle. It is often divided further as heel strike, the support, and toe-off phases.

Swing phase: Consists of the whole time that the foot is within the air. The objectives of the swing phase of gait are to foot clearance over the ground, the forward swing of the limb, and preparation of limb for stance. The swing phase is often weakened into sub-phases. Figure 3 shows the representation of both the phase.

3.2 Calculations

Calculation of average step length is necessary for further calculations. The stride length is the measure of the distance covered while completing one cycle, which is more specific terms can be defined as gait cycle, consider two feet in walking position, let us consider the movement of the left foot first, the left foot is lifted in the air and stepped in front of the right one, now similarly the right foot is lifted and kept in front of left one, now the distance traveled during this motion is called ***stride length***. The ***step length*** is just the first half movement of stride, i.e., the right foot kept in front of the left foot. Table 1 shown below gives information about age, weight, height, and number of steps of different people. This calculation is necessary because of some factors that affect the stride length, and they are age, height, injury, illness, and terrain. For comparison purposes, if two people have the same height but differ in age, the length of stride and steps may be different.

Table 1 Age, weight, height, and number of steps taken by different age category people during their gait

Age	Weight	Height (ft.)	No of steps
17	63.5	5′2″	6 in 11′
21	71.6	5′10″	5 in 11′
48	65.9	5′1″	6 ¼ in 11′
52	71.1	5′5.5″	5 ¾ in 11′
75	52	4′11″	9 ½ in 11′

$$\text{Calculated step length} = \text{Total distance traveled/Number of steps} \tag{1}$$

- For person 1 (age 17) Step length = 11/6 = 1′83″
- For person 2 (age 21) Step length = 11/5 = 2′2″
- For person 3 (age 48) Step length = 116.25 = 1′76″
- For person 4 (age52) Step Length = 11/5.75 = 1′91″
- For person 5 (age 75) Step length = 11/9.5 = 1′15″.

$$\text{Now, the average step length} = \sum\nolimits_{1}^{5} \text{step length/Total observations} \tag{2}$$

$$\text{Average step length} = (1.83 + 2.2 + 1.76 + 1.91 + 1.15)/5 = 1.77 \text{ (calculated)}$$

$$\text{Stride length} \approx 2 \times \text{(Step Length)} \tag{3}$$

$$\text{Average stride length} \approx 2 \times \text{(average step length)} \tag{4}$$

The average stride length of ≈ 3.54

The weight distribution of every part of the body on either side is the same. This implies that the gait of a paralyzed person can be improved by using some wearable devices.

- ***Standing Position***: In this position, the complete weight of the person falls equally on the ground by both the feet. As both the feet impart the same force on the ground, we can call it the total weight of the body. We have considered the subject's weight as 72.6 kg.
- ***Heel Strike***: Let us consider that the right foot is moved first for a gait, and then, the right foot's heel will strike the ground first, making the left foot take more body pressure than the right foot. The exoskeleton is installed on the left foot which is now having more pressure than the right foot, and this pressure will trigger the sensor and send information to the controller and linear actuator as shown in Fig. 4. The total weight of a person is 72.6 kg, and half side body weight is 36.3 kg. The observed weights as the pressure on the left foot increased were 53.5, 50.4, 51.6, 53.8, and 51.5 kg.

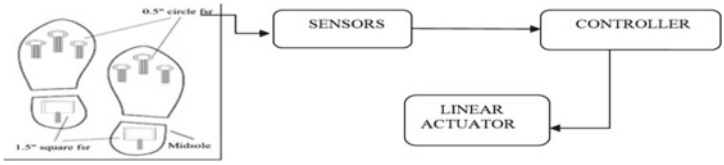


Fig. 4 Positioning of the sensor and connections by blocks

The average from these weights is taken, and that average weight is then used to operate the actuator for limb movement.

The average weight of the heel strike = (53.5 + 50.4 + 51.6 + 53.8 + 51.5)/5 = **52.16** kg.

The parts of the shoe include insole, midsole, and outsole. The insole is the part of the shoe where we place the feet, and the outsole is the part that strikes the ground while walking. The midsole is the part in which sensors will be placed as shown in Fig. 4.

The linear actuator will be functioned by relays, and the operation of walking which is based on the linear actuator will use switching which is done in relays. The specified load capacity and speed of the actuator complete their requirement which will be driven by a lithium-ion battery.

4 Results

The speed of the motor varies with users due to the variation in the weight of different users. The speed of the motor without any load is 20 mm/s, and with its full capacity of load which is 1200 N, it has a speed of 8 mm/s. The plot of force (N) versus resistance (M-Ohm) in Fig. 5 indicates that as the force on the sensor is increasing, correspondingly, the resistance is decreasing as the person keeps his un-paralyzed limb in front of the paralyzed ones; the paralyzed limb having the linear actuator will be activated to perform the motion of gait. The values of force on the x-axis (0.61 N, 1.78 N, 2.35 N, 3.17 N, 4.12 N, 4.36 N, 4.95 N, 5.55 N, 6.6 N) are taken

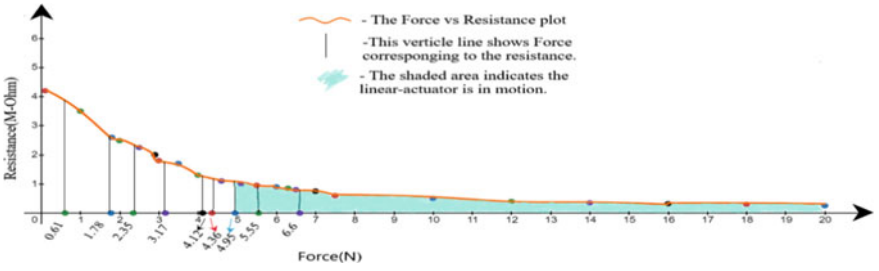


Fig. 5 Observed force versus resistance graph

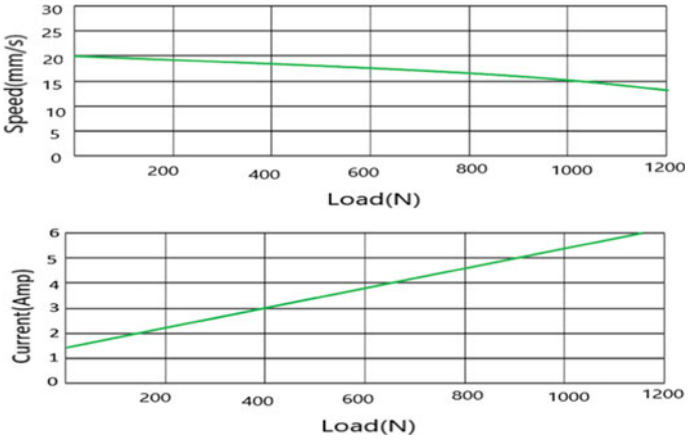


Fig. 6 Speed versus load and current versus load graph

from force-sensing resistor (FSR) while a person was performing the gait. For that person, the toe of the front leg to the heel of another leg makes a total distance of 2.2 feet. The force measured by the FSR was analyzed by considering the left limb is paralyzed.

Here, as the force on the left foot is increasing as well as resistance is decreasing, it makes the linear actuator to function. The shaded part of the graph starts from 4.95 N (actuator will be in motion).

Here, in Fig. 6, the speed versus load graph shows an increasing load from 0 N to 1200 N. The speed is decreased just by 2.5 mm/s from 0 N to 600 N, which ensures less decrement of speed while motioning, whereas, in the current versus load plot, as the load is increasing on the actuator, the current drawn also increases gradually. And Fig. 6 shows the gradual increment of current when the load in newtons has been increased from 0 to 1200 N.

The total weight of the chassis of the exoskeleton is 4.057 kg (as shown in Fig. 7) which also includes the weight of linear actuator which is 1.35 kg. The applied material is given to chassis iron because of its availability and physical properties. Another main advantage of this exoskeleton is that it has a unique design that reduces total manufacturing cost because this exoskeleton is provided for one leg only.

5 Conclusion

The existing models target a wide range of problems such as proper gait, increasing the strength of the person for working in industries, rehabilitation, and many more. Many of these exoskeletons for paralyzed have been developed regardless of the cost and weight of the product which plays a very vital role in developing any system.

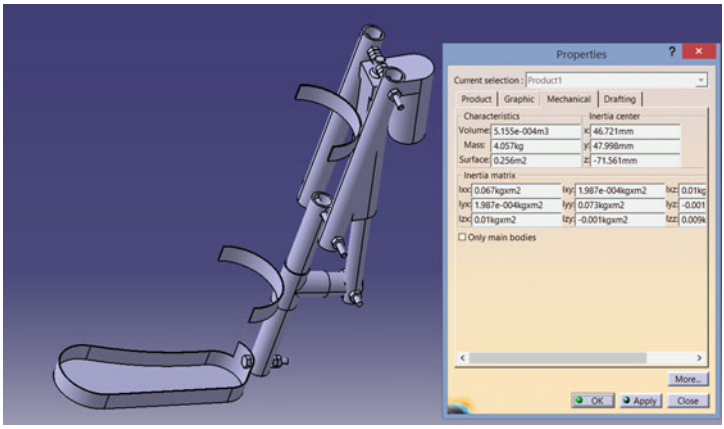


Fig. 7 CAD model with physical properties

The proposed model here accounts for either of the legs, which can be used for rehabilitation, and performing gait if one of the two legs is not paralyzed. By the results gained from the linear actuator and FSR, the exoskeleton will move a distance that is pre-recorded and pre-analyzed from the patients' gait. Belt strap located on the system for tightening the system to limb would be used to adjust according to the individual's foot size which makes the system fully adjustable for different age categories of users. Focusing on the controller, NodeMCU is being used which is very cheap, with good storage capacity (suitable for the exoskeleton), small size, and high onboard processing. The future step includes the complete IOT-based model which can include Bluetooth/Wi-Fi for controlling the exoskeleton, and the NodeMCU has an inbuilt Wi-Fi module, which can help in less complex and more user-friendly systems. For the person who is having both the side of the body paralyzed can also use this system, but this also comes under future scope with the same logic to perform the gait, and the model is designed for a person having gait problems on either side as shown in Fig. 8. The exoskeleton will have a two-sided waking system, and this will also help in the rehabilitation of the patient.

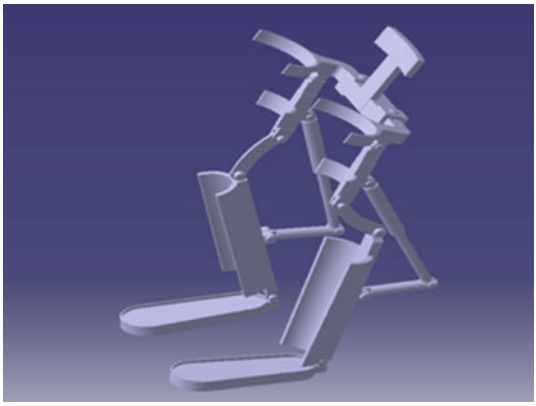


Fig. 8 Proposed design of two-leg exoskeleton

References

1. Kasaoka, K., & Sankai, Y. (2001, October 11). Predictive control estimating operator's intention for stepping-up motion by exo-skeleton type power assist system HAL. In *International Conference on Intelligent Robots and System IEEE* (pp. 1578–1583).
2. Yang, W., et al. (2018). Simulation of exoskeleton ZMP during walking for balance control. In *9th International Conference on Mechanical and Intelligent Manufacturing Technologies IEEE* (pp. 172–176).
3. Zeilig, G., et al. (2012). Safety and tolerance of the ReWalk (TM) exoskeleton suit for ambulation by people with complete spinal cord injury: A pilot study. *Journal of Spinal Cord Medicine, 35*(2), 96–101.
4. Leslie VanHiel PT. "Esko Bionics: Webinar" eksobionics.com 18.04(2).
5. Pirker, W., & Katzenschlager, R. (2017). Gait disorders in adults and the elderly. *Wiener Klinische Wochenschrift, 129,* 81–95.

Detection of Roads in Satellite Images Using Deep Learning Technique

Suvarna G. Kanakaraddi, Ashok K. Chikaraddi, B. L. Pooja, and T. Preeti

Abstract The detection of roads from satellite images is a heated area of research in recent years. Satellite and aerial images are the most important available data sources for map generation and updating available maps. Task of automatically detecting roads is one of the specific cases of this problem. Proposed task is a difficult foresight problem because of occulations, shadows and a huge variety of non-road objects [1]. This research proposes identifying roads by means of a neural network with millions of trainable weights which sees at a much bigger context than the ones used in earlier activation functions and dropout layers. Moreover, real-time image augmentation was verified to improve the accuracy of the model and avoid overfitting. According to experiments, CNN model outperforms all the other tested methodology.

Keywords Satellite images · Convolutional neural networks · Dropout layers · Leaky ReLu · Overfitting · *k*-fold cross-validation · Keras · Kaggle

1 Introduction

Segmentation is a method that is ending up progressively well known for different assignments in computer vision. Generally, the procedure comprises of naming all aspects of an image as per certain criteria. For example, such a method could be utilized for face recognition in photographs or recognition of roads in self-governing driving vehicles.

S. G. Kanakaraddi (✉) · A. K. Chikaraddi · B. L. Pooja · T. Preeti
K L E Technological University, Hubli, Karnataka, India
e-mail: suvarna_gk@kletech.ac.in

A. K. Chikaraddi
e-mail: chikaraddi@kletech.ac.in

B. L. Pooja
e-mail: blpooja95@gmail.com

T. Preeti
e-mail: preeti@kletech.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_44

As of late, the expansion of computing execution, just as the capacity to exploit greatly parallel calculation with GPUs, has prompted the improvement of new AI strategies that can produce images in sensible time. However, image handling has consistently been a difficult assignment, as the data is sorted out in an unequivocal geometrical structure, and in this way calculations ought to think about their morphology. Likewise, the computational expense to process an image does not scale straightly with its size, and this has prompted the advancement of new methods, for example, convolutional neural networks, which cultivates sparse connections and weight sharing so as to decrease the intricacy of the issue.

The aim of the proposed work is to assemble a model that can play out the segmentation of satellite images. In particular, the segmentation comprises of finding which parts of the images are roads, and which parts are foundations (e.g., structures, fields, water). Then the system gives a short diagram of various strategies that can be utilized to take care of this issue, and especially it tends to use convolution neural networks, which speak to the best in class strategy for image characterization.

Recently, deep neural networks have been utilized to create state-of-the-art results about various subfields of computer vision. This classification of calculations can be utilized to extricate generous measure of data from numerous sorts of imagery. The work centers on preparing a neural system to precisely identify roads, by examining a large number of satellite images.

1.1 Road Detection Analysis

The proposed methodology utilizes deep learning model [2], where the framework figures out how to recognize roads from expert-labeled information. Learning methodologies are especially appropriate to the road recognition task since it is an uncommon case of an issue where expert-labeled information is rich. It is difficult to get high resolution aerial images [3] and adjusted guide maps for numerous square kilometers [4]. Learning-based ways to deal with road recognition [5] are not new—a few endeavors at foreseeing whether a specified pixel is road or non-road given features alienated from some context around it have been made.

To start with, small amount of training information is utilized, likely unique images and its ground truth for preparing and testing are commonly acquired by physically labeling an aerial image of every pixel as road or non-road making it infeasible to utilize a ton of training information subsequently, either an exceptionally little setting is utilized to extricate the highlights, or just a couple of highlights are removed from the specific circumstance. At long last, forecasts for every pixel are made freely, disregarding the solid conditions between the road/non-road names for close by pixels. The task expects to construct a model that can perform satellite image segmentation [6]. In particular, the segmentation comprises of recognizing which parts of the images are roads and which parts are background (e.g., structures, fields, water).

2 Related Work

The study of road extraction [7] from high goals multispectral images is carried out where the point of our venture is the extraction of road for the evaluation of a road database [8]. Advances like P-SVM, DSM and SVM classification strategies consolidate the highlights of roads in satellite and protect the solid discriminative capacity of SVM. The proposed framework can concentrate, and string-based parallelism will accelerate the calculation. Yet, the proposed calculation bombs in segregating road structures in profoundly perplexing and jumbled scenes.

As roads have distinctive width and surface material attributes in urban and rustic territories, a measured methodology [9] for road extraction calculation is wanted. The author has assessed the examination to create programmed road extraction calculation in satellite images. It likewise identifies non-exclusive roads from solitary unmanned aerial vehicles (UAV) picture utilizing two noteworthy segments: programmed age of road/non-road seeds and seeded division of road [10].

A vision put methodology [11] for road discovery with respect to printed roads as of now introduced on the MOB-LAB land vehicle to distinguish the road markings quicker by utilizing the geometrical change and morphological preparing has been proposed. It likewise utilizes PAPRICA parallel framework for the continuous road location, which dealing with level roads with painted road markings is fruitful. It arrived at preparing pace of around 17 Hz.

The extraction of specular inherent element from an image is performed and also it utilizes the innovations stereo vision-based 3D road extraction and axis adjustment calculation technique [12] to shift through false identified pixels and to get exact 3D road parameters. Come about location is utilized in a stereo vision-based 3D road parameters extraction calculation. Quick road discovery incorporates shadow and sky expulsions, certainty interim hypothesis and stereo vision location which gives increasingly steady and exact at a diminished computational expense. In any case, disadvantage of vision framework is their reasonableness to enlightenment conditions, for example, shadows, backdrop illumination and low rising sun conditions.

3 Proposed System

The proposed design must be easy enough for an amateur to perceive, understand and detailed enough to describe all the major working modules of the system [13]. The proposed system shown in Fig. 1 comprises loading the dataset of road satellite images which encourages the user to load to the framework. The framework chooses the training road dataset for extraction. When the training image is chosen, it portions satellite images (naming every pixel) by utilizing convolution neural networks. The labeled image is then pre-prepared utilizing the model called convolution neural

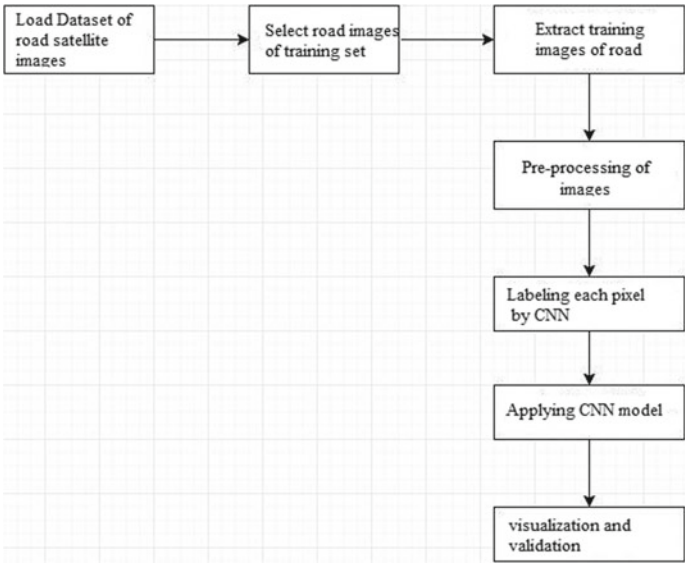


Fig. 1 Description of proposed system

network (CNN) model. From the training results, epoch values are changed over to weights, and these weights help to classify the test set images by utilizing CNN model. Finally, visualization and validation are done for road satellite images by foreseeing the accuracy for the images.

4 Development

This section discusses brief explanation on the portrayal of software system and exhaustive work process. The methodology provides the total examination of each procedure associated with Fig. 2.

4.1 Exploratory Data Analysis

The dataset comprises 100 satellite images of urban territories and their particular ground truth masks, where white pixels speak to roads (closer view) and dark pixels speak to the rest (background). The objective is to group squares of 16 * 16 pixels, taking into account that the label related to each square compares to 1 if the average estimation of the ground truth pixels in that square is more prominent than a limit (0.25), 0 generally. By looking at the training set, it may be seen that the classification undertaking is not inconsequential, as certain roads are secured by trees. Moreover,

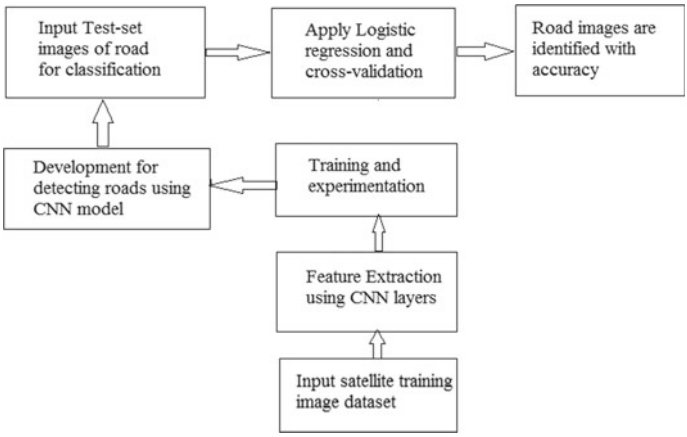


Fig. 2 Block diagram of a system

some black top regions are not marked as road (e.g., parking areas and walkways), and this could conceivably confuse the training model. Figure 3 represents these confusions. Thus, it is pleasing to feel that the classifier should contemplate a kind of setting, for example it should take a gander at close by pixels so as to deduce some data about the square that is being arranged.

Fig. 3 Analysis of dataset

4.2 Feature Extraction Methods

Activation Functions
ReLUs are the standard decision for profound neural networks; however, when a high learning rate is utilized, a few units can stall out and cause the purported dead filters. This issue can be relieved by utilizing a lower learning rate, at the expense of a more drawn out training time (which can be as of now unreasonable). Consequently, a variation known as leaky ReLU has been utilized as the activation function for every intermediate layer, with great outcomes. It is characterized as $f(x) = \max(\alpha x; x)$, with $\alpha \ll 1$, and for our situation α has been equivalent to 0:1. In spite of the fact that this may appear a high worth, a few examinations have demonstrated that higher qualities perform superior to bring down ones, and with this dataset $\alpha = 0{:}1$ has demonstrated viable to avoid dead channels, as appeared in Fig. 4.

Image Augmentation
Since the dataset is little (100 images), an image augmentation procedure has been embraced to basically build its size. In particular, before being provided to the neural system, each training test (i.e., window) is randomly pivoted in ventures of 90 degrees, and it is likewise arbitrarily flipped on a level plane/vertically. This successfully yields an expansion of the dataset size by a multiplicative factor of 8 and has been appeared to incredibly improve the exactness of the model.

Regularization
In spite of the fact that the data augmentation diminishes overfitting, the utilization of dropout layers has been compelling in our model. They have been included after every maximum pooling layer (with $p = 0{:}25$) and furthermore after the completely associated layer (with $p = 0{:}5$). Also, L2 regularization has been utilized for the weights (and not inclinations) of the completely associated and yield layers, with $\lambda = 10^{-6}$. The window size has been experimentally picked so as to consider a setting

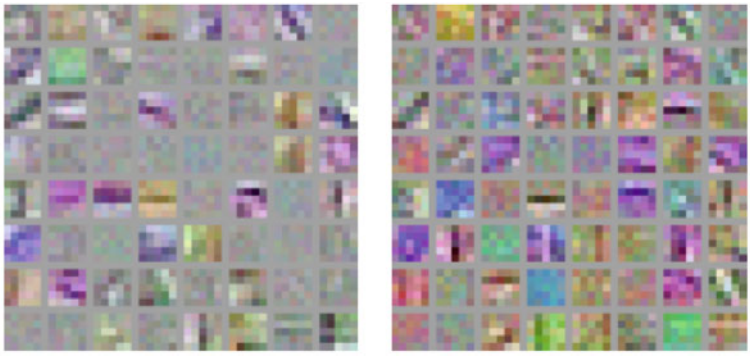


Fig. 4 Visualization of activation functions

Table 1 Layers in neural network

Type	Notes
Input	72 * 72 * 3
Convolution + Leaky ReLU	64, 5 * 5 filters
Max pooling	2 * 2
Dropout	$p = 0{:}25$
Convolution + Leaky ReLU	128, 3 * 3 filters
Max pooling	2 * 2
Dropout	$p = 0{:}25$
Convolution + Leaky ReLU	256, 3 * 3 filters
Max pooling	2 * 2
Dropout	$p = 0{:}25$
Convolution + Leaky ReLU	256, 3 * 3 filters
Max pooling	2 * 2
Dropout	$p = 0{:}25$
Fully connected + Leaky ReLU	128 neurons
Max pooling	$p = 0{:}5$
Dropout	2 neurons

that is enormous enough, taking into account that huge windows are computationally costly. Subsequently, a size of 72 * 72 has demonstrated to be a decent trade-off. Table 1 demonstrates the total structure of the proposed neural system, which is the consequence of different experiments.

4.3 Model Learning

Logistic Regression: Logistic regression is a technique for characterizing information into discrete results. In this module, we present the idea of arrangement, the cost function for calculated relapse and the utilization of strategic relapse to multi-class classification.

Sigmoid Function (Logistic Function) Logistic regression algorithm exploits a linear equation with independent predictors to predict a value. The predicted value can be any place between negative infinity to positive infinity. We need the yield of the calculation to be class variable, i.e., 0-no and 1-yes. In this manner, we are squashing the yield of the direct condition into a scope of [0, 1]. To squash the predicted value somewhere in the range of 0 and 1, we utilize the sigmoid capacity.

$$Z = \theta 0 + \theta 1.x1 + \theta 2.x2 + \ldots \tag{1}$$

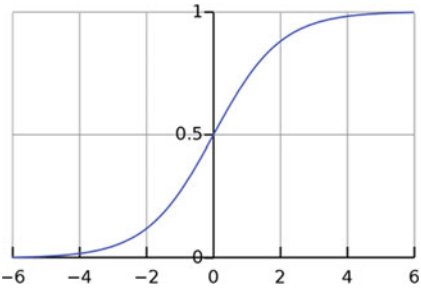


Fig. 5 Sigmoid function graph

$$g(x) = \frac{1}{1+e^{-x}} \qquad h = g(z) = \frac{1}{1+e^{-z}}$$

(a) Sigmoid Function (b) Squashed output-h

To see how sigmoid function squashes the qualities inside the range is imagined in the diagram, Fig. 5, of the sigmoid capacity below: From figure, the sigmoid function moves toward becoming asymptote to $y = 1$ for positive estimations of x and moves toward becoming asymptote to $y = 0$ for negative estimations of x.

4.4 Analysis

The neural network has been actualized and prepared through the Keras library, which is appropriate for prototyping and can utilize either TensorFlow or Theano as back end. One issue with per-pixel classifications is that a stride (i.e., how much the sliding window moves between each training sample) must be picked. Ideally, to ensure that the model is move invariant, the stride ought to be equivalent to 1; nonetheless, this prompts an amazingly huge training set which does not fit in memory, and the condition is declined by the dataset enlargement. As a result, we have actualized a constant preparing set generator that fills in as pursues: At every iteration, the algorithm creates a mini-batch composed of sampled windows from the original training set; all images are changed w.r.t the image augmentation system and their ground truth is determined; at long last, they are provided to the learning algorithm. This technique guarantees that, in the long run, each conceivable zone of the training set will be investigated. Besides, the algorithm is executed on a separate thread, consequently causing no performance hit. For what concerns the conduct at image fringes, reflect limit conditions have been connected, for example the image is reflected along the boundary axis, as appeared in Fig. 6.

The loss function is the softmax straight out cross-entropy, with programmed learning rate modification once the precision arrives at a level for a specific number of emphases. Moreover, the underlying learning rate is set to $\eta = 0{:}001$, and the cluster size is 125 examples.

Fig. 6 Mirror boundary conditions

5 Result Analysis

In light of the outcomes obtained for the relating organizations, road segmentation and detection are demonstrated as follows.

5.1 Validation on Training Set Using CNN Model

To verify the accuracy and exactness of the model, it must be tested with a known validation set. The training dataset has been isolated into two subsets: 75% of the dataset has been utilized to actually train the model (training set), while 25% has been utilized to perform cross-validation (validation set). Utilizing CNN model, training procedure has been done on both sets where each training set images is padded by applying mirror limit conditions. From Table 1 which employs about regularization segment, CNN model gives the quantity of parameters for the given input CNN layers type and yields shape as appeared in Fig. 7.

On our trained model, 146 epochs are finished, and every epoch contains respective validation loss and validation precision as shown in Fig. 8.

5.2 k-*fold Cross-Validation*

The efficient method corresponds to a partial fourfold cross-validation with no long execution cycles. They have been validated with full k-fold cross-validation ($k = 4$) algorithm on the training model to estimate the final accuracy along with their standard deviation. Partial fourfold cross-validation results with accuracy of 0.78528, and

Fig. 7 Parameters with input CNN layers

convolution2d_3 (Convolution2D)	(None, 256, 18, 18)	295168	dropout_2[0][0]
leakyrelu_3 (LeakyReLU)	(None, 256, 18, 18)	0	convolution2d_3[0][0]
maxpooling2d_3 (MaxPooling2D)	(None, 256, 9, 9)	0	leakyrelu_3[0][0]
dropout_3 (Dropout)	(None, 256, 9, 9)	0	maxpooling2d_3[0][0]
convolution2d_4 (Convolution2D)	(None, 256, 9, 9)	590080	dropout_3[0][0]
leakyrelu_4 (LeakyReLU)	(None, 256, 9, 9)	0	convolution2d_4[0][0]
maxpooling2d_4 (MaxPooling2D)	(None, 256, 5, 5)	0	leakyrelu_4[0][0]
dropout_4 (Dropout)	(None, 256, 5, 5)	0	maxpooling2d_4[0][0]
flatten_1 (Flatten)	(None, 6400)	0	dropout_4[0][0]
dense_1 (Dense)	(None, 128)	819328	flatten_1[0][0]

Fig. 8 Epochs with validation loss and accuracy

```
62375/62500 [============================>.] - ETA: 0s - loss: 0.0888 - acc: 0.9656
Epoch 00139: reducing learning rate to 1.5625000742147677e-05.
62500/62500 [==============================] - 123s - loss: 0.0887 - acc: 0.9657
Epoch 141/200
62500/62500 [==============================] - 123s - loss: 0.0877 - acc: 0.9659
Epoch 142/200
62500/62500 [==============================] - 123s - loss: 0.0894 - acc: 0.9652
Epoch 143/200
62500/62500 [==============================] - 123s - loss: 0.0876 - acc: 0.9671
Epoch 144/200
62500/62500 [==============================] - 123s - loss: 0.0877 - acc: 0.9664
Epoch 145/200
62375/62500 [============================>.] - ETA: 0s - loss: 0.0903 - acc: 0.9661
Epoch 00144: reducing learning rate to 7.812500371073838e-06.
62500/62500 [==============================] - 123s - loss: 0.0902 - acc: 0.9661
Epoch 146/200
62500/62500 [==============================] - 123s - loss: 0.0867 - acc: 0.9668
Epoch 00145: early stopping
Training completed
```

full k-fold cross-validation ($k = 4$) results for four images with the accuracy [0.78528, 0.78848, 0.786496, 0.781184] and standard deviation, std = 0.00266787705864 are obtained.

6 Conclusions

The neural network segments the image in a human-like manner when classification is performed on the test set. Specifically, the outcomes appear to be proficient even in border cases, for example, lanes secured with trees or image boundaries. However, the model also tends to produce glitches in certain cases, e.g., diagonal roads. This might be caused by the fact that block classification (in 16 * 16 patches) is not

suitable for these cases, as well as the fact that the majority of the training set is aligned either horizontally or vertically. Based on the convolution neural network with dropout layers, leaky ReLU activates the final cross-validation results, and image augmentation is clearly superior with an accuracy of 78.53%.

References

1. Lu, K., et al. (2015). Unstructured road detection from a single image. In *2015 10th Asian Control Conference (ASCC)*. IEEE.
2. Pooja, B. L., Kanakaraddi, S., & Raikar, M. M. (2018). Sentiment based stock market pridiction. In *International Conference on Computational Techniques, Electronics and Mechnical Systems (CTEMS)* (pp. 12–17).
3. Sujatha, C., & Selvathi, D. (2015). Connected component-based technique for automatic extraction of road centerline in high resolution satellite images. *EURASIP Journal on Image and Video Processing*. https://doi.org/10.1186/s13640-015-0062-9.
4. Christophe, E., & Inglada, J. (2007, September 16–October 19). Robust road extraction for high resolution satellite images. In *ICIP 2007. Proceedings of the International Conference in Image Processing* (pp. 437–440). IEEE.
5. Wang, B., Frémont, V., & Rodríguez, S. A. (2014). Color-based road detection and its evaluation on the KITTI road benchmark. In *2014 IEEE Intelligent Vehicles Symposium Proceedings*. IEEE.
6. Unsalan, C., & Sirmacek, B. (2012). Road network detection using probabilistic and graph theoretical methods. *IEEE Transactions on Geoscience and Remote Sensing, 50*(11), 4441–4453.
7. Zhao, Y., Zhao, J., & Li, Q. (2008). Derivation of respiratory signals from single-lead ECG. In *International Seminar on Future Biomedical Information Engineering, IEEE Computer Society* (pp. 15–18).
8. Jin, X., & Davis, C. H. (2005). An integrated system for automatic road mapping from high-resolution multispectral satellite imagery by information fusion. *Information Fusion*, *6*(4), 257–273.
9. Das, S., Mirnalinee, T. T., & Varghese, K. (2011). Use of salient features for the design of a multistage framework to extract roads from high-resolution multispectral satellite images. *IEEE Transactions on Geoscience and Remote Sensing, 49*(10), 3906–3931.
10. Huang, Z., Zhang, J., Wang, L., & Xu, F. (2012). A feature fusion method for road line extraction from remote sensing image. In *IEEE International Geoscience and Remote Sensing Symposium*, (p. 13133956).
11. Zhou, H., et al. (2016). On detecting road regions in a single UAV image. *IEEE Transactions on Intelligent Transportation Systems, 18*(7), 1713–1722.
12. Broggi, A. (1995). Robust real-time lane and road detection in critical shadow conditions. In *Proceedings of International Symposium on Computer Vision-ISCV*. IEEE.
13. Easa, S. M., Dong, H., & Li, J. (2007). Use of satellite imagery for establishing road horizontal alignments. *Journal of Surveying Engineering, 133,* 29–35.

Quantification of Multimillion Offers in 'Next-of-Kin' Unsolicited Bulk Emails

Jatinderkumar R. Saini and Sachin Naik

Abstract Unsolicited Bulk Email (UBE) still continues to successfully pass through the keyword-based and Artificial Intelligence (AI)-based filters due to the smart use of 'word-salad', 'slang' and 'leet' by the spammers to dilute as well as pollute the email text. Many research works account for the wastage of bandwidth and time of users as well as loss of money due to victimization to 'next-of-kin' type of UBE. This is the first formal attempt to present the quantified results on the calculations of averages for transaction amount offered, offered share of total amount, duration of transaction completion and total amount mentioned in such UBE. A peep is also provided on the identification of most frequently used first-names and surnames by the spammers. It is emphasized that this type of UBE does not ask for bank details and asks for money directly, unlike the many other types of UBE often confused with this type. On the side lines, the paper presents an interplay of scam and spam while providing useful statistical information for design of more robust spam email filters and blockers. A corpus of more than 1500 specific and only 'next-of-kin' type of spam emails has been used. It has been found that the average, respectively, for amount and share offered by the spammers is 10.64 million USD and 39.43% of the total amount. Only 5.57, 2.49 and 1.86% of the total number of spam emails mention the offered share, duration of transaction completion and total amount, respectively. The average transaction duration was 9 days while the average of total amount was found to be 134 million USD. 'Aisha' and 'Kabore' were the most frequent first-name and surname respectively.

Keywords Cry-for-help · Email · Next-of-kin · Scam · Spam · Unsolicited bulk email (UBE)

J. R. Saini (✉) · S. Naik
Symbiosis Institute of Computer Studies and Research, Symbiosis International Deemed University, Pune, India
e-mail: saini_expert@yahoo.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_45

1 Introduction

Unsolicited Bulk Email (UBE) is a type of email which is neither solicited nor expected by the recipient. It is a type of spam and also called Unsolicited Commercial Email (UCE) [1]. Email being a cheap and anonymous mode of communication, the said type of email is sent in bulk and across the international boundaries. Many such emails are sent by spammers in hope of successfully victimizing the recipient. Generally, such emails land in our 'Bulk', 'Spam', 'Junk' or similar types of folders [2]. Many a times, though, these emails also land in the Inbox. The relative comprehension of the definition of spam itself poses a challenge for its classification. A mail related to some offer for traveling may be spam for person not interested in it while of great interest to someone very fond of traveling. Had it been just for interest or disinterest, the things would not have been that serious. The problem with spam is that it is not just time-wasting, annoying but also attempts to lure an innocent person in the trap of fake and cheating offers.

The UBE emails are tricky and intend to cheat the recipient. They typically claim of being originated from a person who has a big amount of money to donate, transfer or invest in the recipient's country. The scammers call the amount to be offered to the so-called beneficiary for the help by the later variously as compensation, donation, charity or reward. Isacenkova et al. [3] have called it payoff while Newman [4] calls it 'pay a little, get a lot' scam. The reasons for doing this are generally presented as being a childless widow or on death-bed or in exile. The spammer claims that before initiating the transfer of millions of United States Dollars (USD) in the email recipient's account needs to complete some formalities for which the recipient should pay. It is this trap that many innocent people fall prey to and lose their hard earned money. Generally, out of fear of social embarrassment, these types of cases never surface up. Typically, the scammers try to exploit the element of greed among the human beings while using their spam email as a bait to victimize a prospective innocent recipient of the email. There is no specific strategy, except the brute force approach, with the scammers to identify a specific target email ID to victimize. Their approach typically makes them hunt in the dark for possible victimization. The current work is an attempt to find the threshold amount offered by the scammer to exploit greed in human beings. There are also elements of religion, region, relationship, politics, geography, tragedy, disease, accident, death and donation or charity or all of these that are used by the spammers to create a gloomy picture for generating pathos and pity among the readers for their possible victimization.

The most appropriate approach to fight the threat of spam is to understand the spam, the underlying scam and hence their interplay along with the psychology of the spammer as well as the prospective victim. This is the motivation behind this paper which is a genuine attempt in this direction. There are many interesting results reported in the paper. It is claimed that this is first formal attempt of its kind in world where the quantified values from various 'next-of-kin' emails are reported. The results presented here are also claimed to provide much needed additional strength to the otherwise typical keyword-based algorithms for spam filtering and blocking.

Generally these types of filters fail due to a very smart use of 'leet', 'slang' and 'word-salad' [5] by spammers. The concept of using a large text portion of normal text along with the text of spam email is called word-salad and is implemented by the spammer to dilute the keyword-based filters. Using a name which is similar to the actual name and is interpreted well by a normal human being but not so by a machine is an example of usage of slang and leet. These are also the reasons of inefficiency of many Artificial Intelligence (AI) based algorithms as they function the best when fed with proper training data but misbehave when the data itself is diluted or polluted and hence this paper aims at presenting many quantified and statistical values for 'next-of-kin' types of spam emails where the spammer request the recipient to pose as a next-of-kin of some multi-millionaire deceased foreigner.

2 Literature Review

Though a lot of work has been done for classification of emails into spam and non-spam categories, the work for statistical analysis of spam emails is limited. Kulkarni and Saini [6] attempted to identify whether email has attachments or not. A research work with a limited corpus size of only 300 documents stated that June is the favorite month of the spammers while Friday happens to be the favorite week-day [7]. Bibi et al. [8] presented analysis of non-ham mail scanning by use of Machine Learning algorithms. Venkatraman et al. [9] used the concept of semantic similarity and presented the categorization of UBE for Internet of Things. Al-Rawashdeh et al. [10] and Sanghani and Kotecha [11] used modified feature selection approaches for detection of email spam. Mansourbeigi [12] as well as Elshoush and Dinar [13] used stochastic methods for categorization of UBE. Saini and Desai [14] analyzed the spam emails for identification of non-dictionary words specifically in the emails that claimed to advertise the products for various diseases while they identified the top words from the spam emails which projected to award lottery winnings [15].

Kulkarni et al. [16] analyzed the impact of the header part of the email on the categorization of spam emails while they [17] discussed role of Bag of Words on the categorization of non-spam emails. Rader and Munasinghe [18] discussed about the emails which are misdirected. Ho et al. [19] discussed about an innovative 'meet in the middle' methodology for system meant for the identification of non-ham emails. Wei et al. [20] have attempted to correlate the Internet Protocol addresses used in the spam emails for the purposes of cyber forensics. Li et al. [21] analyzed the spam emails with focus on spread of spam emails as well as viruses. Lehrfeld et al. [22] interestingly claimed that 15% of email users tend to click on the hyperlinks in the emails again and again. Lumezanu and Feamster [23] attempted to correlate spam which was common for both Twitter as well as email messages. Fette et al. [24] concluded that the behavior of sending spam email messages by the scammers is evolutionary in nature.

Isacenkova et al. [3] have presented an analysis of phone numbers used in the scam emails. Newman [4], Atkins and Huang [25], and Cummins [26] claim that

such scammers have also started to use social engineering for making the attacks more precise. Nigerian Scams Scam-watch [27] claims that for the month of February 2020 itself, as per the reported values, victims have lost 3291 USD. They further report that 60% of the victims were females while, irrespective of gender, maximum number of victims were from the age group of 45–64 years. Lin et al. [28] reported that older females are most susceptible. Leonhardt [29] reports that these scams cost a victim at an average of 2133 USD. He further adds that the US residents lost an amount of over 0.7 USD million in the year 2018 alone. Longe et al. [30] provided a black-list of various links, mail addresses, and words that are used by the spammers. Ibrahim [31] argued that a combination of various social and cultural parameters contributed to Nigerian children becoming criminals. He presented the results based on his personal interaction with their parents.

Hence, researchers in the past have focused well on spam emails [32], the classification of spam email messages [33] as well as analysis of its categorization [34]. They have worked for mining the spam emails for analyzing them based on use of email ID, gender, age group, usage of Internet protocol, social engineering, and hyperlinks in email messages, to name a few. But none of them have attempted to figure out the average of compensation amount offered, committed transaction share as well as the duration during which the transaction will be completed. In order to further better understand the spam email industry, this paper analyzes spam email messages from these and more dimensions. On the sidelines, the paper also presented the state of the art of spam email classification as well as analysis.

3 Methodology

In this section, a detailed description of the methodology followed for achieving the objectives is presented. A corpus consisting of a total of 1562 spam emails of type 'next-of-kin' were collected. In order to pre-process as well as process this corpus, Java and C programming languages were used. In order to fine tune the results and for specific screening, some part of processing was also done manually. The rest of this section is organized by inclusion of sub-sections on calculations of averages for transaction amount offered, offered share of total amount, duration of transaction completion and total amount mentioned in the spam email. Also, included is a description of steps followed for identification of most frequently used first-names and surnames by the spammers. The last sub-section describes the process followed for identification of various pieces of information sought by the spammers from the recipient of the email.

3.1 Calculation of Offered Average Transaction Amount or Average Compensation Amount

This paper uses the word compensation to represent this entire set of synonymously used words including 'payoff', 'proposal amount' and 'deposit'. The words like 'reward' have not been used intentionally as the same creates an ambiguous impression of the 'award-winning' mails and the mails claiming lottery winnings. In order to find the average compensation amount offered by the scammers, all the numeric quantities were extracted from the corpus. These quantities comprising mainly of date of spam email, zip codes, house numbers, share of total amount offered, total amount, year of accident/death/plane-crash and duration during which the transaction will be completed were further filtered to find just the numeric amounts of compensations offered. This compensation amount had further possibilities like total amount claimed to be handled by the scammer, the amount offered for compensation, and the amount and number of installments to be paid by the scammer to the prospective victim.

Another difference with the compensation amount was that it was found to be in United States Dollars (USD), Pounds Sterling and Euros. The individual average amount in all of these currencies as well as the average of averages converted in common denomination of USD was also calculated. The rates prevalent on 3rd April, 2020 were taken into consideration for the necessary conversions. 1 lb Sterling equivalent to 1.23 USD and 1 Euro equivalent to 1.08 USD was taken into consideration. No compensation amount in other currencies like Yen, Dinar, Rupees, Baht and Yuan, to name a few, was found. Instead of using 'Pounds Sterling', the spammers prefer to use 'Pounds' as the currency. The compensation amount mentioned in the spam emails could also be called the transaction amount as this was the amount that the scammer claims to transact with the prospective victim. This amount is different from the total amount, higher than the transaction amount which is part of the total amount, mentioned in a few emails, and discussed in Sect. 3.4.

3.2 Calculation of Average Share of Offered Compensation

While calculating the average compensation offered by the scammers, it was found that all mails contain some amount offered to the prospective victim. In fact, this amount itself is the bait. However, during the calculation of average share of offered compensation it was found that only 87 out of a total of 1562 UBEs contained the numeric value for the share of the compensation amount. It is notable that this value is not straight-forward to be calculated as the search for '%' sign in the corpus yielded 833 UBEs out of a total of 1562. This is because the spammers tried to insist that the transaction is '100% risk free'. Though the most commonly shared offer by the scammers was for 40% for the prospective victim, this figure also had values of 30, 45 and 50%. The average of the available values was calculated for the share as

offered compensation. It is noteworthy that this is the amount of part of the total amount mentioned in the scam. It is only this amount of share that is offered by the criminal as the compensation.

3.3 Calculation of Average Transaction Time Mentioned in the UBE

Very few mails mentioned the duration during which the entire transaction will be completed. The transaction here means the time duration during which the criminal claims to transfer the promised amount to the so-called beneficiary's account. Only 39 mails were found to mention this duration of transaction. This data was used for the calculation of the average transaction time.

3.4 Calculation of Average of Total Amounts Mentioned in the UBE

Most of the spam emails did not mention about the total amount possessed by the criminal. In case it was mentioned, it was also mentioned that how much of this amount will be transacted by the criminal. Only 29 of 1562 spam emails mentioned the total amount and the amount to be transacted differently. For the rest of the mails only the actual amount to be transacted was mentioned with a few mails also mentioning that the criminal and victim will have some share of this total amount to be transacted. The average amount was calculated from the mails containing the total amount and transaction amount separately.

3.5 Identification of Most Commonly Used First-Name and Surname in the UBE

All the 1562 spam emails mentioned the name of the sender of the mail. Though this was the name of the criminal, it was neither true nor trustworthy. A list of all the names mentioned in all the emails was prepared and it was found that the list of the unique names was slightly more than the number of spam emails itself. The list was larger than the number of spam emails itself because there were many spam emails which contained multiple names like 'I am ≪name_1≫, daughter/son/widow/wife/mother/etc. of ≪name_2≫'. It is noteworthy that when the unique names were found from the list of the names, it was found that the list contained 1571 unique names from 1562 spam emails. The frequency of these names was also calculated and the most commonly used names were identified.

3.6 *Identification of Atomic Information Pieces Solicited by Scammers*

Each mail was individually scanned to retrieve the pieces of information solicited by the scammers. It is notable that it is this section of the spam email that is of second most relevance to the scammers. This paper advocates that the first thing of relevance for the scammer is getting the mail recipient fall prey to the offered award. Different scammers use different terms for getting this information. Only atomic pieces of information were considered for further processing for sake of clarity as well as better comprehension. The processed atomic information pieces from the retrieved ones were further clubbed using the concept of 'synset' which is the set of synonyms for a given term. Hence, the words like 'Profession' and 'Occupation' used differently by scammers were clubbed together and considered as a single term only. Similarly, {'Position in Office', 'Designation'}, {'Full name', 'Complete name'} and {'Postal Address', 'Contact Address', 'Correspondence Address'} were sets of other couplets or triplets and terms treated synonymously using synset. In the context of present discussion, we term, for instance, {'Position in Office', 'Designation'} is a synset. We have not made use of any existing English WorldNet for this.

4 Results and Findings

This section presents the results, findings as well as discussion and analysis thereof. A random snapshot of the various entities derived from the spam emails of the scammers and used for the current work is presented in Table 1. The value of 'Not provided' in

Table 1 A random snapshot of various entities derived from the scam emails

S. No.	Transaction amount (in millions)	Total amount (in millions)	Currency	Offered share	Transaction completion duration (in days)
1	15.5	50	USD	Not provided	Not provided
2	2	Not provided	USD	40	Not provided
3	8.5	Not provided	Euro	Not provided	Not provided
4	30	200	USD	50	Not provided
5	1.5	Not provided	USD	Not provided	Not provided
6	15.5	Not provided	Pounds	Not provided	Not provided
7	17.8	Not provided	USD	40	Not provided
8	5.8	Not provided	Euro	Not provided	Not provided
9	3.5	Not provided	Pounds	Not provided	Not provided
…	…	…	…	…	…
1562	22.3	Not provided	USD	Not provided	10

Table 1 indicates that the respective value was not provided by the scammer in the spam email.

1. The average compensation amounts offered by the criminals to the prospective victims, in USD, Pounds Sterling and Euros was found to be 13.67 million, 9.5 million and 6.11 million respectively. When these amounts were converted into common denomination of USD, the amount was respectively 13.67 million USD (the same), 11.67 million USD (converted from Pounds Sterling to USD) and 6.59 million USD (converted from Euros to USD). The average of these three averages yielded an amount of 10.64 million USD.
2. Only 5.57% of cry-for-help spam emails contain the value of share of the total amount mentioned in the spam email and in the form of compensation offered by the criminal to the prospective victim. It was found that the average offered share of the total amount mentioned in the spam email was 39.43%. It was also found that the lowest offered share was of 30% while the highest offered share was of 50%.
3. Only 2.49% of the total number of spam emails mentioned the duration of time during which the criminal claimed to complete the transaction (s) of transferring the offered compensation to the beneficiary's (actually a prospective victim) account. The average time duration was 9 days. The maximum time duration was found to be 14 days while the lowest time duration was found to be 3 days or 72 h.
4. Only 1.86% of the total number of spam emails mentioned the total amount and transaction amount separately. The average of total amount was found to be 134 million USD. The minimum and maximum total amount was 50 million USD and 218 million USD respectively. It is remarkable to mention here that the average of transaction amount is already mentioned as part of point no. 1 of the Results section. It is also remarkable that the average total amount of 134 million USD mentioned in the emails is much higher than the average transaction amount of 10.64 million USD mentioned in the emails.
5. The most commonly used first-name of the sender of the spam emails was found to be 'Aisha', along with its variations with surname, like 'Aisha Al-Qaddafi', 'Aisha El Gaddafi', 'Aisha Gaddafi', and 'Aisha Muammar Gaddafi'. The most commonly used surname was found to be 'Kabore' along with its variations like 'Kabor', 'Kaabore', 'Kaboree', and 'Kaborre'. It is noteworthy to mention that 'Kabore' is a famous surname from a country in West Africa called The Republic of Burkina Faso and the nearby area. A snapshot of first 10 first-names and first 10 surnames, not necessarily correlated with each other, both sorted in ascending order and irrespective of their frequency, is presented in Table 2. It is noteworthy that with a limited corpus size of 300 documents only, it was reported by the researchers using Part-of-Speech (POS) approach and Stanford Parser that 'Savimbi' is the most frequently used name by the senders of the spam emails [7].

Table 2 A random snapshot of lists of first-names and surnames

S. No.	First-name	Surname
1	Abdel	Abbas
2	Abdul	Abbott
3	Abogado	Alphonso
4	Agbor	Al-Qaddafi
5	Ahmad	Altangerel
6	Ahmed	Ander
7	Aisha	Aziz
8	Ali	Bachir
9	Amira	Badza
10	Ana	Bailey

6. Very interestingly, it was found that there is no consistency in the use of various terms used by the spammers for retrieving the same information from the prospective victims. They use different terms like 'Complete name' and 'Full name' to convey the same thing. Table 3 gives the list of unique atomic pieces of information solicited by the scammers. Table 3 presents this information in alphabetically sorted manner. The terms spammer and scammer have been used interchangeably in this paper. Similarly, the description is in context of the 'prospective victim' or the 'email recipient', which are the two terms used interchangeably in this paper.
   The ID solicited by the spammers could be the Passport Number, Social Security Number, Aadhaar Number, Driving License, and so on. It is notable that some spammers seek the Passport Number specifically separately too. When the spammer seeks only the 'email address' or the 'phone number', then it is based on the discretion of mail recipient to interpret it as either official one or the personal one, though there are many emails where the spammer specifically seeks 'private email ID' or the 'private phone number'. Some spammers seek name of 'residence country', 'country in which the compensation amount is to be received' and 'country of origin' separately. Though one of these, particularly the last one, could be treated as an indication of citizenship of the recipient, surprisingly, no spam email was found to seek this information explicitly, at least by using the word 'citizenship'.
   Further, though the spammers seek more than a dozen information pieces from the mail recipient, they do not use the word 'citizenship' explicitly. These unique pieces of information 'mainly' include Name, Age, Sex, Phone Number, Fax Number, Occupation, Country, Position in Office, Passport Number, Email ID, City, Address, Marital Status, and Profession. The Name, Age, Phone, Occupation, and Country were the highest sought pieces of information by the spammers. This was followed by Sex and Address, which in turn were followed by the rest pieces of information including Fax and Passport Number.

Table 3 List of atomic fields of information solicited by scammers from the recipients of the spam emails

S. No.	Atomic field of information solicited by scammer
1	Age
2	City
3	Copy of ID or IDs for identification
4	Country
5	Date of birth
6	Email
7	Fax number
8	Full names
9	Gender
10	Home address
11	Marital status
12	Mobile No.
13	Occupation
14	Office address
15	Passport number
16	Phone number
17	Picture
18	Position in office
19	Postal Address
20	Private email address
21	Private phone number
22	Receiving country
23	Residence country

7. One very important result is that none of the spammers were found to ask for the Bank Account Number or the other Bank details of the recipient of the email. This is important to discuss because of at least two reasons: (a) the mails considered for current research are not 'lottery winning' emails though they are 'phishing emails' (b) the mails considered here are only first tier mails. The current paper treats all those mails considered in this research as first tier while the second tier mails are beyond the scope of current discussion. When the recipient receives the first tier mail, responds back to it and further gets another mail from the spammer, the second mail received from the spammer is treated as the mail belonging to second tier.

5 Conclusion, Implications and Future Work

It is concluded that the spammers use only USD, Pounds Sterling and Euros as the currencies for the offered compensation amounts to the prospective victims. The average offered compensation amount, irrespective of currency is more than 10 million USD. It is further concluded that on an average, the criminals offer an approximate 40% of share, of the total amount mentioned in the spam email, as the compensation to the prospective victim. It was further found that the criminals claimed to complete the transfer of promised compensation amount to the victim's account within 9 days on an average. It is further concluded that less than 2% spam emails from scammers mention the transaction amount and the total amount separately. The average total amount, which is different from the average transaction amount of slightly more than 10 million USD, is 134 million USD. The most commonly used first-name and surname of senders is 'Aisha' and 'Kabore' respectively. It was further observed that different spammers use different terms to seek same piece of information from the recipient of the email. Though labels like 'Name', 'Age', 'Phone', and 'Country' remained the most sought pieces of information by the spammers, they do not directly seek the information about the bank details and the citizenship of the recipient of the spam email. This is also the reason why these 'next-of-kin' type of spam emails should not be categorized 'phishing emails' which typically try getting the bank details of the victim and instead of money directly.

Though the results, discussion and conclusion presented here are best reported for the set of more than 1500 scam emails considered here, the results may be considered generic in nature. With many results being 'first reporting' in the field, this paper is believed to have an important ripple effect among the future as well as contemporary researchers. The results presented here could also be used to enhance the various spam filtering algorithms used by different mail service providers which generally do not consider these details in lack of their availability. Also, the statistical results presented here are promising to make the spam filters robust against their 'only keyword-based' counterparts which often fail in wake of use of word-salad, slang and 'leet' by the spammers and scammers. The correlation of results presented here with the similar results derived from other spam email categories will be the future work to further extend the presented research. This is believed to provide better platform to understand spam, scam, their interplay as well as the spamming behavior of the scammers.

References

1. Lam, H., & Yeung, D. (2007). A learning approach to spam detection based on social networks. In *Fourth Conference on Email and Anti-Spam (CEAS-2007).*
2. Saini, J. R., & Desai, A. A. (2009). *Self Learning Taxonomical Classification System using Vector Space Document Analysis Model for Web Text Mining in UBE, PhD Thesis accepted by Department of Computer Science.* Surat, India: Veer Narmad South Gujarat University.

3. Isacenkova, J., Thonnard, O., Costin, A. et al. (2014). Inside the scam jungle: a closer look at 419 scam email operations. *EURASIP Journal on Information Security*, *4*. https://doi.org/10.1186/1687-417X-2014-4.
4. Newman, L. H. (2020). Nigerian email scammers are more effective than ever. https://www.wired.com/story/nigerian-email-scammers-more-effective-than-ever/. Last accessed 02 Apr 2020.
5. Uemura, T., Ikeda, D., & Arimura, H. (2008). Unsupervised spam detection by document complexity estimation. In: J. F. Jean-Fran, M. R. Berthold & T. Horváth (Eds.), *Discovery science. DS 2008. Lecture notes in computer science*, (Vol 5255). Berlin, Heidelberg: Springer.
6. Kulkarni, P., & Saini, J. R. (2020). Identification of minimal email header features for discovery of email attachments. *International Journal of Advanced Science and Technology, 29*(6), 6242–6247.
7. Saini, J. R. (2012). Identification and analysis of most frequently occurring significant proper nouns in 419 Nigerian scams. *National Journal of Computer Science and Technology, 4*(1), 1–4.
8. Bibi, A., Latif, R., Khalid, S., Ahmed, W., Shabir, R. A., & Shahryar, T. (2020). Spam mail scanning using machine learning algorithm. *Journal of Computers, 15*(2), 73–84.
9. Venkatraman, S., Surendiran, B., & Raj, P. A. K. (2020). Spam e-mail classification for the internet of things environment using semantic similarity approach. *The Journal of Supercomputing, 76*(2), 756–776.
10. Al-Rawashdeh, G., Mamat, R., & Rahim, N. H. B. A. (2019). Hybrid water cycle optimization algorithm with simulated annealing for spam e-mail detection. *IEEE Access, 7,* 143721–143734.
11. Sanghani, G., & Kotecha, K. (2019). Incremental personalized E-mail spam filter using novel TFDCR feature selection with dynamic feature update. *Expert Systems with Applications, Elsevier, 115,* 287–299.
12. Mansourbeigi, S. M. H. (2019). Stochastic methods to find maximum likelihood for spam e-mail classification. *AINA Workshops, 2019,* 623–632.
13. Elshoush, H. T., & Dinar, E. A. (2019). Using adaboost and stochastic gradient descent (sgd) algorithms with R and orange software for filtering e-mail spam. *CEEC, 2019,* 41–46.
14. Saini, J. R., & Desai, A. A. (2011). Identification of non-lexicon non-slang unigrams in body-enhancement medicinal UBE. *World Academy of Science, Engineering and Technology, 80*(286), 1631–1637.
15. Saini, J. R., & Desai, A. A. (2011). Identification of most frequently occurring lexis in winnings-announcing unsolicited bulk e-mails. *World Academy of Science, Engineering and Technology, 4*(3), 172–176.
16. Kulkarni, P., Saini, J. R., & Acharya, H. (2020). Effect of header-based features on accuracy of classifiers for spam email classification. *International Journal of Advanced Computer Science and Applications, 11*(3), 396–401. https://doi.org/10.14569/IJACSA.2020.0110350.
17. Kulkarni, P., & Saini, J. R. (2020). A creation of bag of words using map reduce technique for ham email classification. *International Journal of Advanced Science and Technology, 29*(5), 715–721.
18. Rader, E., & Munasinghe, A. (2019). Wait, do i know this person?: Understanding misdirected email. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI'19)*, (Vol. 290, pp. 1–13). New York: ACM. doi:https://doi.org/10.1145/3290605.3300520.
19. Ho, P., Kim, H., & Kim, S. (2014). Application of sim-hash algorithm and big data analysis in spam email detection system. In *Proceedings of the 2014 Conference on Research in Adaptive and Convergent Systems (RACS '14)*, (pp. 242–246). New York: ACM. doi:https://doi.org/10.1145/2663761.2664221.
20. Wei, C., Sprague, A., Warner, G., & Skjellum, A. (2008). Mining spam email to identify common origins for forensic application. In *Proceedings of the 2008 ACM symposium on Applied computing (SAC '08)*, (pp. 1433–1437). New York: ACM. doi:https://doi.org/10.1145/1363686.1364019.

21. Li, W., Hershkop, S., & Stolfo, S. J. (2004). Email archive analysis through graphical visualization. In *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security (VizSEC/DMSEC '04)*, (pp. 128–132). New York: ACM. doi:https://doi.org/10.1145/1029208.1029229.
22. Lehrfeld, M., Ogle, A., Franklin, B., & Dangler, J. (2014). Dimensioning SPAM: An in depth examination of why users click on deceptive emails. *Journal of Computer Science College, 30*(2), 213–219.
23. Lumezanu, C., & Feamster, N. (2012). Observing common spam in twitter and email. In *2012 Internet Measurement Conference (IMC'12)*, (pp. 461–466). New York: ACM. doi:https://doi.org/10.1145/2398776.2398824.
24. Fette, I., Sadeh, N., & Tomasic, A. (2007). Learning to detect phishing emails. In *Proceedings of the 16th international conference on World Wide Web (WWW '07)*, (pp. 649–656). New York: ACM. doi:https://doi.org/10.1145/1242572.1242660.
25. Atkins, B., & Huang, W. (2013). A study of social engineering in online frauds. *Open Journal of Social Sciences, 1*(3), 23–32.
26. Cummins, E. (2020). The Nigerian prince scam is still fooling people. Here's why, https://www.popsci.com/story/technology/nigerian-prince-scam-social-engineering/. Last accessed 02 Apr 2020.
27. Nigerian Scams (2020). https://www.scamwatch.gov.au/types-of-scams/unexpected-money/nigerian-scams. Last accessed 02 Apr 2020.
28. Lin, T., Capecci, D. E., Ellis, D. M., Rocha, H. A., Dommaraju, S., Oliveira, D. S., & Ebner, N. C. (2019). Susceptibility to spear-phishing emails: effects of internet user demographics and email content. *ACM Transactions on Computer—Human Interact, 26*(5). doi:https://doi.org/10.1145/3336141.
29. Leonhardt, M. (2019). 'Nigerian Prince' email scams still rake in over $700,000 a year-here's how to protect yourself. https://www.cnbc.com/2019/04/18/nigerian-prince-scams-still-rake-in-over-700000-dollars-a-year.html. Last accessed 07 Apr 2020.
30. Longe, O. B., Abayomi-Alli, A., Shaib, I. O. L., & Longe, F. A. (2009). Enhanced content analysis of fraudulent Nigeria electronic mails using e-STAT. In *2009 2nd International Conference on Adaptive Science & Technology (ICAST-2009)*, (pp. 238–243). Ghana: Acra.
31. Ibrahim, S. (2016). Causes of socioeconomic cybercrime in Nigeria. In *2016 IEEE International Conference on Cybercrime and Computer Forensic (ICCCF)*, (pp. 1–9). Vancouver: BC.
32. Saini, J. R., & Desai, A. A. (2010). A survey of classifications of unsolicited bulk emails. *National Journal of Computer Science and Technology, 2*(1), 33–39.
33. Saini, J. R., & Desai, A. A. (2010). A supervised machine learning approach with re-training for unstructured document classification in UBE. *INFOCOMP Journal of Computer Science, 9*(3), 30–41.
34. Saini, J. R., & Desai, A. A. (2010). Analysis of classifications of unsolicited bulk emails. *World Academy of Science, Engineering and Technology, 4*(2), 91–95.

Prevailing Standards in Requirement-Based Test Case Prioritization: An Overview

Omdev Dahiya and Kamna Solanki

Abstract Testing the software is a time-consuming and costly process. Testing teams are often constrained to end their testing endeavors soon, owing to time and budget constraints. This can lead to issues with the satisfaction of the customer and the quality of the product. Test case prioritization (TCP) techniques have shown to improve the viability of regression-testing activities. Due to these, the fault detection rate is increased, which allows testers to discover faults early in the software product. In the testing phase, the necessity of requirements information has been very widely documented by the requirements engineering community. Most of the techniques for regression testing rely upon the code information of the software. Including the requirements information to the existing testing, techniques will help the testers in identifying the source of the defects more quickly and validating the product corresponding to its requirements. This paper has focused on finding the current state of the art of requirements-based test case prioritization techniques in the field of software testing.

Keywords Regression testing · Requirement-based test case prioritization · Software testing

1 Introduction

The introduction of software-enabled devices has led to a revolution in the form of digital transformation. They have made the lives of humans very easy, as almost all the tasks can be easily accomplished by their assistance [1]. With such reliability on the software product, there comes a responsibility for developing a reliable product. For ensuring it, the developed product is tested for its effectiveness [2]. A large

O. Dahiya (✉) · K. Solanki
University Institute of Engineering and Technology, Maharshi Dayanand University Rohtak, Haryana, India
e-mail: Omdahiya21792@gmail.com

K. Solanki
e-mail: Kamna.mdurohtak@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_46

number of test cases are designed to achieve this task of testing it thoroughly but, it is practically not possible to execute all of them due to limited time, budget and expert availability, etc. [3]. Owing to these reasons, there is a need for a reduction in the test case number. The test case prioritization approach of regression testing helps in reducing the test case number. It does not delete test cases, but it arranges them according to priority, where higher priority test cases are executed first, and then the ones with lower priority are executed [4].

2 Objectives of This Study

The objective of this paper is to document the progression made by the researchers in the requirement-based test case prioritization techniques. For it, work done by the numerous researchers is selected and comprehensively studied. It will lay the foundation for this study. This study will summarize the research findings of eminent researchers. The readers will be benefitted who wants to work in this field, and this paper will serve as a comprehensive source of information on this topic.

3 Research Methodology

For selecting the relevant studies related to this topic, we have followed a few steps. Firstly, a string for searching the relevant articles is defined, which is "requirements-based test case prioritization." Secondly, the database is selected for searching the relevant papers. This study has used Google Scholar as a platform because almost all the reputed publishers such as IEEE Xplore, Scopus, Springer, ACM Digital Library, other reputed publications, and conference proceedings are available at a single platform. Then, the selection of studies is performed from it. Available studies are refined by following various steps, such as title-based exclusion, abstract-based exclusion and conclusion-based exclusion. When a study has passed all of these in the form of relevancy to our topic, then it is studied thoroughly and included in our study. Overall, from the searched results, we can finalize 21 studies to be included in this study, which have fulfilled the desired parameters.

4 Results and Analysis

This section presents the research work of numerous researchers who have worked toward proposing requirement-based test case prioritization techniques for regression testing. Observations about a particular study have been discussed in detail.

Srikanth and Williams, "Requirements-based test case prioritization" (2002).
The authors have worked toward making the test case prioritization process cost-effective and at the same time, earlier detection of faults. They have told that this is their basic and primary study, and they are planning to extend their work using industrial data [5].

Srikanth et al., "System test case prioritization of new and regression test cases" (2005).
The authors have proposed a technique that uses several factors for the prioritization of test cases. They have demonstrated the efficiency of their technique in terms of the increased rate of fault detection. They have concluded that customer priority is one of the essential factors in prioritization based on requirements [6].

Zhang et al., "Test case prioritization based on varying testing requirement priorities and test case costs" (2007).
The authors have performed this study to address quality testing issues and to enhance the satisfaction of the user. This study has assumed that requirements and costs are varying while developing a project [7].

Srivastava et al., "Test case prioritization based on requirements and risk factors" (2008).
The authors in this study have prioritized the requirements instead of just prioritizing the test cases. They have stressed that testers must decide that the test case corresponds to the requirements being targeted [8].

Kavitha et al., "Requirement-based test case prioritization" (2010).
The authors have proposed an algorithm to use requirements information for prioritizing the test cases. They have demonstrated that the fault detection rate is increased. They have even used industry-level projects to show the effectiveness of their technique [9].

Salem and Hassan, "Requirement-based test case generation and prioritization" (2010).
The authors have performed this study, mapping the test cases to their corresponding requirements to be tested. This will eventually lead to complete coverage of the requirements framed for the project and will overall lead to customer satisfaction [10].

Maia et al., "Applying search-based techniques for requirements-based test case prioritization" (2010).
The authors have demonstrated the usage of search-based techniques for prioritizing the test cases. They have performed it by considering the characteristics of the requirements [11].

Salehie et al., "Prioritizing requirements-based regression test cases: A goal-driven practice" (2011).
The researchers have focused on two issues in this study, firstly on cost, quality and efforts for a project and using metrics that do not require code for evaluating. They have shown how requirement-based test case prioritization is efficient [12].

Arafeen and Do, "Test case prioritization using requirements-based clustering" (2013).
The authors have worked toward investigating whether the usage of information about requirements for test case prioritization is beneficial or not. Their results proved that it is very beneficial as by using that information, testers can quickly identify the error-prone modules [13].

Muthusamy and Seetharaman, "A new effective test case prioritization for regression testing based on prioritization algorithm" (2014).
The authors have presented an algorithm that prioritizes the test cases based upon requirements and several other factors. They have shown the effectiveness of their algorithm by comparing it with other techniques [14].

Ma et al., "Test case prioritization based on requirement correlations" (2016).
The authors have proposed a technique to readjust the order of requirements that are related to faults. They have performed experiments, and their results show high test effectiveness [15].

Srikanth et al., "Requirements-based test prioritization using risk factors: An industrial study" (2016).
The authors have conducted this study as an extension of their prior work. They have used industrial data to observe the efficiency of their algorithm under different conditions. They have concluded that its usage in the industry will be fruitful in terms of scheduled delivery of the projects and in addressing budget-related issues [16].

Wang and Zeng, "History-based dynamic test case prioritization for requirement properties in regression testing" (2016).
The authors have conducted this study following a history-based approach for requirements priorities. They have used industrial projects and have shown promising results in terms of enhanced fault detection rate [17].

Thakurta, "Understanding requirement prioritization artifacts: a systematic mapping study" (2017).
The authors have conducted this study to lay out the factors that influence the prioritization of requirements. The authors emphasized its necessity, as not all requirements can be met with available resources and time [18].

Alzaqebah et al., "Whale optimization algorithm for requirements prioritization" (2018).
The authors have proposed a nature-inspired technique for prioritizing the requirements. They have tested and compared their technique with other existing approaches and have proved the supremacy of their technique [19].

Masadeh et al., "Grey Wolf algorithm for requirements prioritization" (2018).
The authors have proposed an algorithm based on nature-inspired techniques for prioritization of requirements. They have shown the efficiency of their technique by comparing them with other approaches [20].

Hujainah et al., "Software requirements prioritisation: a systematic literature review on significance, stakeholders, techniques, and challenges" (2018).
The authors have reviewed the existing literature and concluded that how prioritization of requirements has led to developing a quality product with specified constraints. They have documented the shortcomings in the existing techniques so that much better can be proposed by the researchers [21].

Ambreen et al., "Empirical research in requirements engineering: trends and opportunities" (2018).
The authors have conducted this study to document the current trends in the field of requirements prioritization. They have selected studies from numerous databases and have shown the trends in this field as well as future opportunities in it [22].

Yaseen et al., "Requirements Prioritization and using Iteration Model for Successful Implementation of Requirements" (2019).
The authors have proposed a technique for ranking the requirements. They have shown how project estimation time is reduced, and the timely delivery of a product is ensured. Their study has shown a new dimension in prioritizing the requirements [23].

Jarzabek et al., "Inferring Hints for Defect Fixing Order from Requirements-to-test-case Mappings" (2020).
The authors in their research work have provided a method to resolve the defects with manageable efforts, which are found while performing regression testing. They have defined how requirements should be mapped to corresponding test cases for test effectiveness [24].

Table 1 Summative list of researches included in this study

S. No.	Authors name	Publication year
1.	Srikanth and Williams	2002
2.	Srikanth et al.	2005
3.	Zhang et al.	2007
4.	Srivastava et al.	2008
5.	Kavitha et al.	2010
6.	Salem and Hassan	2010
7.	Maia et al.	2010
8.	Salehie et al.	2011
9.	Arafeen and Do	2013
10.	Muthusamy and Seetharaman	2014
11.	Ma et al.	2016
12.	Srikanth et al.	2016
13.	Wang and Zeng	2016
14.	Thakurta	2017
15.	Alzaqebah et al.	2018
16.	Masadeh et al.	2018
17.	Hujainah et al.	2018
18.	Ambreen et al.	2018
19.	Yaseen et al.	2019
20.	Jarzabek et al.	2020
21.	Nayak et al.	2020

Nayak et al., "An Improved Approach to Enhance the Test Case Prioritization Efficiency" (2020).
The authors have proposed a technique to make the test case prioritization more efficient. They have shown how their technique reduces the requirement of test cases as compared to other approaches [25].

Table 1 provides a summative list of researchers, along with the publication year of their study.

5 Conclusion and Future Scope

In this paper, current researches and existing literature in the field of requirement-based test case prioritization algorithms proposed by numerous researchers have been documented. By this, the current state of the art in this domain has been presented. As in the software testing sphere, a significant role is played by the requirements. The developers and end-users define the definition of the requirements prioritization. Various studies have highlighted that mapping the test cases to their corresponding

requirements to be tested will eventually lead to complete coverage of the requirements framed for the project and will overall lead to customer satisfaction. This study will serve as a platform for the source of information to the aspiring researchers in this field to go through the existing literature and will aid in proposing new and effective techniques. The future direction of this study will try to perform experimental comparisons and put the results more systematically.

References

1. Khatibsyarbini, M., Isa, M. A., Jawawi, D. N., & Tumeng, R. (2018). Test case prioritization approaches in regression testing: A systematic literature review. *Information and Software Technology, 93,* 74–93.
2. Yoo, S., & Harman, M. (2012). Regression testing minimization, selection, and prioritization: A survey. *Software Testing, Verification, and Reliability, 22*(2), 67–120.
3. Dahiya, O., & Solanki, K. (2018). A systematic literature study of regression test case prioritization approaches. *International Journal of Engineering & Technology, 7*(4), 2184–2191.
4. Mukherjee, R., & Patnaik, K. S. (2018). A survey on different approaches for software test case prioritization. *Journal of King Saud University—Computer and Information Sciences.*
5. Srikanth, H., & Williams, L. (2002). Requirements-based test case prioritization. *IEEE Transactions on Software Engineering*, 28.
6. Srikanth, H., Williams, L., & Osborne, J., (2005). System test case prioritization of new and regression test cases. In *2005 International Symposium on Empirical Software Engineering, 2005,* (pp. 1–10). IEEE.
7. Zhang, X., Nie, C., Xu, B., & Qu, B. (2007). Test case prioritization based on varying testing requirement priorities and test case costs. In *Seventh International Conference on Quality Software (QSIC 2007)*, (pp. 15–24). IEEE.
8. Srivastva, P. R., Kumar, K., & Raghurama, G. (2008). Test case prioritization based on requirements and risk factors. *ACM SIGSOFT Software Engineering Notes, 33*(4), 1–5.
9. Kavitha, R., Kavitha, V. R., & Kumar, N. S. (2010). Requirement based test case prioritization. In *2010 International Conference on Communication Control and Computing Technologies*, (pp. 826–829). IEEE.
10. Salem, Y. I., & Hassan, R. (2010). Requirement-based test case generation and prioritization. In *2010 International Computer Engineering Conference (ICENCO)*, (pp. 152–157). IEEE.
11. Maia, C. L. B., De Freitas, F. G., & De Souza, J. T. (2010). Applying search-based techniques for requirements-based test case prioritization. In *I WOES-CBSoft*, (pp. 24–31).
12. Salehie, M., Li, S., Tahvildari, L., Dara, R., Li, S. & Moore, M. (2011). Prioritizing requirements-based regression test cases: A goal-driven practice. In *2011 15th European Conference on Software Maintenance and Reengineering*, (pp. 329–332). IEEE.
13. Arafeen, M. J., & Do, H. (2013). Test case prioritization using requirements-based clustering. In *2013 IEEE Sixth International Conference on Software Testing, Verification and Validation*, (pp. 312–321). IEEE.
14. Muthusamy, T., & Seetharaman, K. (2014). A new effective test case prioritization for regression testing based on prioritization algorithm. *International Journal of Applied Information Systems (IJAIS), 6*(7), 21–26.
15. Ma, T., Zeng, H., & Wang, X. (2016). Test case prioritization based on requirement correlations. *In 2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD)*, (pp. 419–424). IEEE.
16. Srikanth, H., Hettiarachchi, C., & Do, H. (2016). Requirements based test prioritization using risk factors: An industrial study. *Information and Software Technology, 69,* 71–83.

17. Wang, X., & Zeng, H. (2016). History-based dynamic test case prioritization for requirement properties in regression testing. In *Proceedings of the International Workshop on Continuous Software Evolution and Delivery*, (pp. 41–47).
18. Thakurta, R. (2017). Understanding requirement prioritization artifacts: a systematic mapping study. *Requirements Engineering, 22*(4), 491–526.
19. Alzaqebah, A., Masadeh, R., & Hudaib, A. (2018). Whale optimization algorithm for requirements prioritization. In *2018 9th International Conference on Information and Communication Systems (ICICS)*, (pp. 84–89). IEEE.
20. Masadeh, R., Alzaqebah, A., Hudaib, A., & Rahman, A. A. (2018). Grey Wolf algorithm for requirements prioritization. *Modern Applied Science, 12*(2), 54.
21. Hujainah, F., Bakar, R. B. A., Abdulgabber, M. A., & Zamli, K. Z. (2018). Software requirements prioritisation: A systematic literature review on significance, stakeholders, techniques, and challenges. *IEEE Access, 6,* 71497–71523.
22. Ambreen, T., Ikram, N., Usman, M., & Niazi, M. (2018). Empirical research in requirements engineering: Trends and opportunities. *Requirements Engineering, 23*(1), 63–95.
23. Yaseen, M., Ibrahim, N., & Mustapha, A. (2019). Requirements Prioritization and using Iteration Model for Successful Implementation of Requirements. *International Journal of Advanced in Computational Science Applications, 10*(1), 121–127.
24. Jarzabek, S., Liszewski, K., & Boldak, C. (2020). Inferring hints for defect fixing order from requirements-to-test-case mappings. In *Integrating Research and Practice in Software Engineering*, (pp. 43–51), Cham: Springer.
25. Nayak, S., Kumar, C., Tripathi, S., & Majumdar, N. (2020). An improved approach to enhance the test case prioritization efficiency. In *Proceedings of ICETIT 2019*, (pp. 1119–1128). Cham: Springer.

Text-Based Handwritten Recognition Through an Image Using Recurrent Neural Network

D. Kalyani and P. Vijay Kumar

Abstract Neural computers' computational operations are employed in a strategic way that is entirely disparate from that of standard conventional computers. Neural computers are primarily computational train tutored there by if given a definite precise original state (data input), they whatsoever classify the input data should be rolled into one of the well known categorized classes or convicts that raw input data to yield into a specific worthwhile advantageous property is improved. In this system, we pass small patches of handwritten images to convolution neural networks, and it recognizes the text from the given input images. Neural computers implement data parallelism. The main objective of this paper is to propose the design of an expert knowledge-based neural network system for handwritten recognition that can effectively recognize the text from the given input image using recurrent neural network approach. This approach can be used as an optimal solution tool for handwritten recognition as it provides faster classification with more accuracy.

Keywords Neural networks · Handwritten character · Image segmentation · Recurrent neural networks

1 Introduction

1.1 Neural Networks

A neural network in short NN is a robotic model which embodies many associated layers of neurons. A neuron having a mathematical function results in a number, which is also known as an activation. Neuron links are assigned weights, which characterize the strength of the associated neuron. Input data is taken and then nurtured into the first layer of neural network, which activates input each neuron

D. Kalyani (✉) · P. Vijay Kumar
VNR Vignana Jyothi Institute of Engineering & Technology, Bachupally, Hyderabad, India
e-mail: kalyani_d@vnrvjiet.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_47

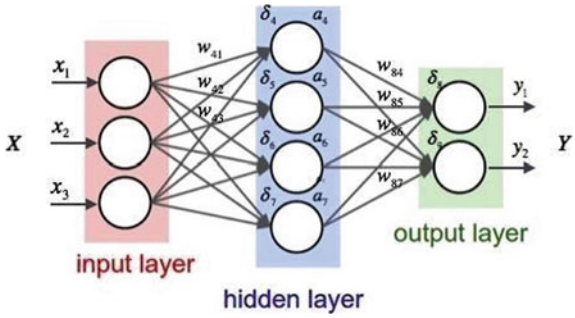


Fig. 1 Neural network

to certain extent. Based on the activation function and weights, the network identifies the neurons that are to be activated from the following layers. This mechanism is known as feedforward process, which is persisted until the output neurons are enabled. The framework of a neural network has control on the data that it should work with to improve the final achieved output. The subsequent Fig. 1 represents a three-layered simple neural network consisting of input, hidden and output layers.

We can also represent the NN mathematically [1], i.e., in further formal sense such as a mathematical function, which maps an impression (or a numerical matrix), size of $M\ W \times H$ to a character string represented as $(c_1, c_2, \ldots, c_n)$ with length in the midst of 0 and L. In this, the input is conceded at character level; thereby, texts that are not equipped in the practice data can be realized too.

$$NN : M \rightarrow (c_1, c_2, \ldots, c_n); 0 \leq n \leq L \tag{1}$$

1.2 Types of Neural Networks

(i) Convolutional Neural Networks or CNN:

Convolutional neural network known as CNN uses three basic factors for the implementation of classification and recognition problem. They are separated into two parts known as firstly, a feature learning part, and another one is a classification part with either one or multiple layers in each. Feature learning part is normally performed by merging convolution tiers and pooling layers [2, 3]. Basically, CNN typically consists of three types of convolution, pooling and fully connected layers as elementary building blocks. The first two layers are used to perform feature extraction, whereas the last layer maps the features that are extracted into the final classification output. An example for CNN is shown in Fig. 2.

A convolution layer is the first layer of the CNN design which performs linear and nonlinear operations for feature descent. Kernel is a small set of array numbers used as a filter. This filter is applied across an array of numbers, which is the input, called

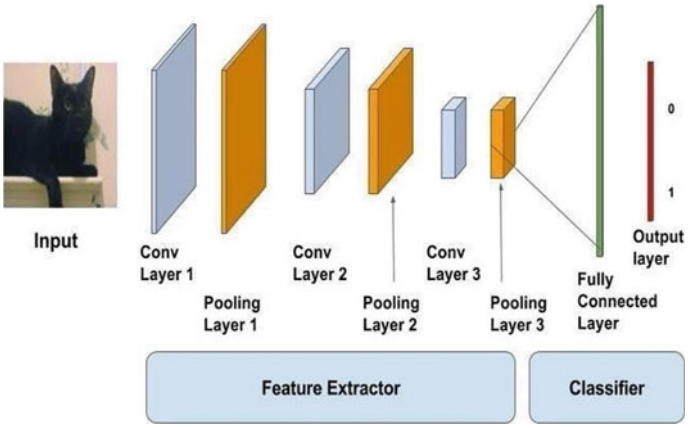


Fig. 2 Convolutional neural network

a tensor. Product is calculated element-wise at each location of the tensor linking the input tensor and every other element of the kernel. Later, output is summed at the corresponding locations of the output tensor, known as a feature map. One of the most importnat characteristic of tensors' is to have multiple kernels repeatedly for the formation of arbitrary number of feature maps [4].

Pooling layer: A pooling layer is used to reduce the dimensional magnitude of the feature maps which helps to introduce translation invariance for distortions caused and for small shifts.

But padding, filter size and stride are considered as hyper parameters in pooling operations that are similar to convolutional operations. Max pooling is one of the forms of pooling operation that is most popularly known. This max pooling is used to extract patches from the input feature maps and is also used to output each patch with its maximum value. Then, it discards all the other values. A filter which is of size 2×2 is sufficient in this max pooling. This down samples the feature maps in-plane dimensionality by a factor of 2.

Fully connected layer: Fully connected layers are connected to every output layer by a learnable weight. Here, in this layer, the outputs of the feature maps which are obtained in pooling or convolutional layer are transformed into an array of numbers or vectors which are of one-dimension(1D) connected to one or more of these fully connected layers followed by a nonlinear function, such as the rectified linear unit (ReLU).

(ii) Recurrent Neural Networks (RNN):

RNN is one of the crucial types of neural networks. In RNN, it takes the input from the output of the previous step and does further processing. In the traditional mostly used neural networks, all the data inputs and resultant output are not related to each other as they are independent from one another. So that, after predicting a word or character,

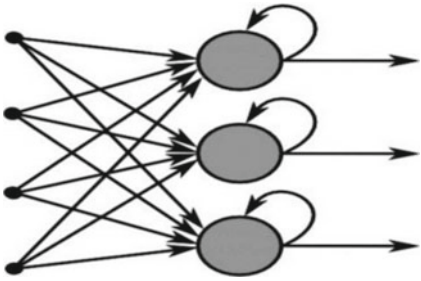


Fig. 3 Recurrent neural network

it does not remember anything in the next prediction. If we have the previous words or characters that are predicted, can be stored or remembered, then it will be easy to predict the next characters or words based on the previous ones. To solve this problem and get the next predictions before only based on the previous predictions, RNN has come into the existence, to solve this problem using the concept of a new layer called hidden layer. This hidden layer is the most predominant attribute of RNN, as shown in Fig. 3, which is also used to store and remember about the sequence it has predicted.

RNN has a special term called as memory, which is used to remember all the information about what has happened and calculated before. It executes the similar kind of task on all the hidden and input data layers which uses the same parameters for every input to produce the results or outputs. This helps to reduce complexity of parameters.

(iii) LSTM (Long Short Memory Networks):

LSTM is abbreviated for long short memory networks, which are of a special noteworthy type of known RNNs. They are used to learn over long-term dependencies. This makes RNN capable of remembering the collected inputs over a long time. They are capable of remembering by storing their information in a memory like the memory stored in a computer. LSTMs after storing their information, they can write the data, read the data and delete the data. This is how LSTM improves the RNN with an efficient computational complexity of O (1) and bridge within the time interval.

2 Literature Survey

Various existing methodologies introduced by various authors that use for the identification of text from the handwritten input images have been summarized here:

Recognition of Handwritten Character using NN
This system is developed to recognize the characters in the given scanned documents and also to investigate the effects on changing models of artificial NN [5].

Recognition for Handwritten English Letters
These days distinctive systems are utilized for character recognition. Archive confirmation, advanced library, perusing bank store slips, perusing the postal locations, removing the data from checks, information section, applications for mastercards, medical coverage, credits, tax documents and so forth are the application territories of computerized report handling. This system research works for the handwritten English letters recognition [1, 2]. Manually written letters are hard to perceive because of the various human penmanship style, variety in point, size and state of the letters. Different methodologies of the written by hand character acknowledgment are talked about in this paper alongside their execution.

Devanagari Recognition of Handwritten Characters through Artificial NN and Segmentation
Manually written Devanagari characters are relatively more mind-boggling for acknowledgment than comparing English characters because of numerous conceivable varieties altogether, number, bearing and state of integral strokes. The main mechanism for unplugged Devanagari handwritten string recognition by proposed feedforward artificial neural network-based segmentation [6] of these characters at first into line, word and finally to characters.

Diagonal Feature Extraction-Based Handwritten Alphabets Recognition With NN
Another neural network feature extraction strategy known as diagonal-based feature extraction [7, 8] for extracting the hand written text which contains 26 alphabets with varying dissimilar hand strokes . We can extend it to five hundred and seventy alphabetical characters handwritten in numerous styles are in use for experiment purpose. This has given larger amounts of the acknowledgment precision contrasted with the system utilizing the horizontal and vertical strategies [9] for the feature extraction process.

A Technical Construct on Character Geometry in Recognition of Character
A geometry-based technique is used in this system for the feature extraction process which is applicable for based on segmentation word recognition methodologies. The geometric properties of character border on basic line types [10] extracted in turn form the character skeleton in this system.

Handwritten Character Recognition using a Robust Gradient Feature Extraction Method
The basic piece of technique for any of the recognition systems is to depict the input templet by methods which are more powerful in discriminating pattern classes [9]. This system ensures an effort made to obtain a recognition accuracy of 94% in recognition of the English handwritten characters. Due to its ease of the use toward logical simplicity and high recognition rate, gradient features ability have numerous real-time applications.

Recognition of Hindi Characters using Backpropagation Neural Networks
Recognition of manually written Hindi characters recognition system using a neural networks is a troublesome errand in light of the fact that these characters are written in different sizes, thickness with varying bent and cursive styles, so that they could be of various kind of font organization and measuring parameters. This system proposed an approach to recognize Hindi characters focusing on insensitive missing data in four stages (1) Scanning input data; (2) Pre-processing for noise reduction, also banalizations for normalization and thinning; (3) Feature extraction incorporates separating some valuable data out of the diminished picture as a feature vector; (4) Recognition: A back-propagation neural network is utilized to attain 93% as an average accuracy of recognition for this system.

Devanagari Character Recognition Using Neural Networks
In this system, another robust NN approach has been proposed to recognize real-time automatic offline characters using multilayer perception with the hidden layer. There is also a significant extension work which can be applied for any Indian language like Hindi, Canada, Tamil, Bangla, Malayalam, and so forth. However Devanagari is a language [6, 11] for which barely any work is detectable particularly using binary character recognition mechanism and the same is stored in the form of matrix of size 'n × n'.

Handwritten Devanagari using NN
In the present computerized time, the most vital thing is to manage the advanced records, and associations utilizing manually written archives for the putting away of their data can utilize transcribed character [12] acknowledgment to change over this data into advanced organization. Written by hand of Devanagari characters are increasingly troublesome for acknowledgment because of the nearness of conjunct characters in a header line and similitude in the states of various characters. According to this, each individual input data character or numerical picture is first partitioned into 'n' same sized grids or areas [13, 14]. It registers the normal separation of all pixels present in every one of the networks concerning picture centroid just as zonal areas centroid character component vector of the size '2 × n.'

3 Proposed Methodology

The proposed methodology for implementing the handwritten character recognition is to implement the image processing techniques. To recognize the characters, we take the handwritten images as an input, process the characters in the input images and train neural network algorithms using convolutional and recurrent neural networks in order to get high accuracy for the recognition of characters. This proposed design can be implemented in Python using Jupiter ide, in every procedure such as image pre-processing, processing and post-processing. The main required packages for this project are OpenCV, NumPy and TensorFlow. The key objective of this design is to

study and implement a methodology that can be able to recognize the text in images, which are hand written in English language only.

In our proposed system, the given input image is injected to train the CNN layers in order to extract admissible features from the given input image while performing three basic operations. They are listed as—(1). The convolution operation applies a kernel filter (5 × 5) to the input in the first two layers and 3 × 3 in the last three layers; (2). A nonlinear rectified linear unit (ReLU) function is put in; (3). A pooling layer is applied, which encapsulates the image territory regions, and output yields a retrenched interpretation of the given input. In this process, feature maps are used to diminish the height of the image by half in each and every layer resulting in its size reduced to 32 × 256.

3.1 Selection of Input Image

In this strategy, the information is taken as the handwritten images. Neural networks consist of simple tiny elements influenced by biological nervous system neurons which operate in parallel. We can accustom the neural network by reconciling the weighted values over all available connections in between the elements in order to perform a particular mathematically defined function. In fact, neural networks are trained, in such a way that each particular input data gives a specific determined target output. That means all NNs are trained by comparing the output of each phase with the target. This process continues until the output matches with the given target. Moreover, many input/targets pairs are used as a part of supervised learning to train a network as per the requirement. The input is taken from the handwritten images of different persons and the images available on the Internet.

We consider a gray value-based image which is of size 128 × 32. Usually, in real time, these images that are taken from the dataset do not have the exact size as that of our dimensions as mentioned. If that is the case, then we resize the image without any sort of distortion until it attains its dimensions reach to either the height to 32 or a width to 128. In the next step, we copy the same image into a target image with dimensions 128 × 32. Later, the original input values of the image are normalized so that it can simplify the operational task for the neural network. Training of the images is fed into different multiple layers using convolutional and recurrent neural networks in order to extract different features. Finally, we record certain parameters like training time, accuracy, etc., to find the effectiveness of our proposed neural network system. Outputs are obtained after every stage in pre-processing, segmentation and recognition.

3.2 Design Procedure of Proposed Approach

To recognize the handwritten characters, we need to perform analysis in three different stages. They are:

- Pre-processing stage—Gray scaling and noise removal.
- Processing stage—Image segmentation, classification and recognition.
- Post-processing stage—Connectionist temporal classification.

In this proposed system, 256 features are contained in every feature sequence per time step. As the popular implementation of RNN's LSTM is used, it can be moved through longer distances to propagate information and robust training characteristics are provided when compared to RNN's output feature sequence and finally mapped into a matrix which is of size 32 × 80. In our system, we have taken the real-time dataset which contains 79 differently abled characters. For this purpose, we have made 80 entries for each step of 32 time stamps. Now, the output of CNN layer is passed as input to the RNN layers for further processing. Many features are thus extracted from the CNN layers, and some more features that are not extracted from CNN can be identified and extracted in RNN layers.

3.3 Training Through RNN

1. Single step of input at once is to be provided through the chosen network system.
2. Enumerate its current state using a current input state and previous output state.
3. The current state represented as 'h_t' becomes (h_t-1) in the next time stamp.
4. Step 3 can be repeated as many time steps as required to join the outputs from all the previous states as per the problem that has been defined.
5. The final current state is used after completion of all the time steps to compute the output.
6. The output obtained from Step 5 is now verified by comparison with the actual output required; i.e., the target output for a match otherwise an error is resulted.
7. In case of error, it is later back-propagated to RNN for updating of weights with a new set of values, and hence, the network, i.e., RNN, is trained.

4 Implementation Results

4.1 Pre-processing Stage

Pre-processing stage is crucial as it normalizes and then removes variations like noise using normalization, filtering techniques for skew detection, noise removal and

slant correction. In this proposed model, we have implemented two pre-processing techniques: gray scaling, re-sizing and normalization.

Gray Scaling: Grayscale pictures can be the consequence of estimating the power of light at every pixel. Conversion into gray scale is always not wished for. When converting into gray scale, you decrease the amount of picture information, as well as lose data. For some applications, picture handling applications shading is critical, and changing over to grayscale can decline comes about.

In our scenario, we take handwritten image as an input and then transformed into a grayscale image. As the input that is taken itself is in the form of grayscale, so then no need of grayscale conversion but when the input image is in 'rgb' format then the conversion to gray scale is required. This step reduces the processing and computational speed and also the memory space is also optimized.

Re-sizing: In image re-sizing, the true boundary of the strokes is determined in terms of cropping redundant background area and the given image is resized to a standard size. Center of contour is mostly preferable for the samples than the center of gravity as it is more resilient to noise.

Normalization: The main component of the pre-processing stage includes cleaning, line detection, skew correction, character size normalization and slant and slope removal in normalization phase without affecting the identity of the word.

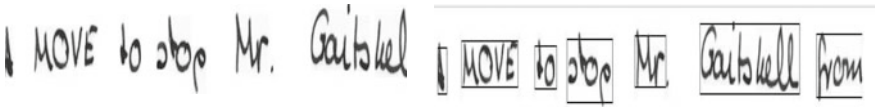
4.2 Processing Stage

The major process of this proposed design involves image segmentation, classification and recognition.

Image segmentation: In the image segmentation phase itself major processing takes place, the input is taken as image which consists of a set of words and characters are broken into sub-images to have a strapping correlation with areas or objects of the image. Segmentation is an important stage as the recognition rate directly depends upon the extent to which one can separate lines, words and characters. Image segmentation is the process in which every pixel is assigned a label such that the pixels with the same label can share visual characteristics. The goal of segmentation is to simplify the image into some other form that is more flexible to analyze as shown in Fig. 4a and b.

Steps of the Algorithm:

1. Creating kernel filter and Applying filter kernel.
2. Find connected components using OpenCV.
3. Append components to result.
4. Append bounding box and image of word to result list.
5. Return list of words, sorted by x-coordinate.



a **b**

Fig. 4 **a** Before segmentation. **b** After segmentation

Fig. 5 Implementation code snippet

```
def wordSegmentation(img, kernelSize=25, sigma=11, theta=7, minArea=0):
    kernel = createKernel(kernelSize, sigma, theta)
    imgFiltered = cv2.filter2D(img, -1, kernel, borderType=cv2.BORDER_REPLICATE).astype(np.uint8)
    (_, imgThres) = cv2.threshold(imgFiltered, 0, 255, cv2.THRESH_BINARY+cv2.THRESH_OTSU)
    imgThres = 255 - imgThres

    if cv2.__version__.startswith('3.'):
        (_, components, _) = cv2.findContours(imgThres, cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_SIMPLE)
    else:
        (components, _) = cv2.findContours(imgThres, cv2.RETR_LIST, cv2.CHAIN_APPROX_SIMPLE)

    res = []
    for c in components:

        if cv2.contourArea(c) < minArea:
            continue

        currBox = cv2.boundingRect(c)
        (x, y, w, h) = currBox
        currImg = img[y:y+h, x:x+w]
        res.append((currBox, currImg))

    return sorted(res, key=lambda entry:entry[0][0])
```

Implementation code is provided in Fig. 5.

4.3 Post-processing Stage

Connectionist Temporal Classification (CTC): CTC is added as output layer where the image gets transformed into a conditional probability distribution. Because RNNs are less efficient to image segmentation and detecting cursive characters, CTC allows RNN output matrix to sequence directly the ground truth text and finally computes the loss value. The ground truth text as well the recognized text must be at most 32 characters. Observed results for beam search are shown in Fig. 6.

Fig. 6 Beam search for input image

5 Conclusion and Future Work

The language scripts in the entire world would have varying handwritten characters that leads to a difficult situation to recognize the existence of similarities or dissimilarities in shapes for multiple characters. Our evolution algorithm pre-processes the input image with proper RNN-based training technique. Finally the output results are non-negligibly efficient when compared with existing statistical features extraction methods of neural networks. In this paper, we bring up effective ways to train a CNNRNN hybrid architecture using synthetic data for stronger recognition results of English characters. Our work further can be extended to convert the fax and newspapers into text format and for reading postal address, etc.

We can use multiple NNs to recognize words, sentences or paragraphs. For classification and character segmentation, a greedy search-based approach is the most likely suitable solution. The use of LSTM can be replaced by two-dimensional LSTM. We can also improve the optimizer which indirectly improves the accuracy rate.

References

1. Sharma, N., et al. (2013). Recognition of handwritten English letters: A Review. *International Journal of Engineering and Innovative Technology, 2*(7), 318–321.
2. Kaur, M., & Kumar, S. (2012). A recognition system for handwritten gurmukhi characters. *International Journal of Engineering Research and Technology*.

3. Tschopp, F. Efficient convolutional neural networks for pixel wise classification on heterogeneous hardware systems.
4. Mehfuz, S., & Katiyar, G. Intelligent systems for off-line handwritten character recognition: A review. *International Journal of Emerging Technology and Advanced Engineering*.
5. Borde, S., Shah, E., Rawat, P., & Patil, V. (2012). Fuzzy based handwritten character recognition system. *International Journal of Engineering Research and Applications*.
6. Dongare, S. A., Kshirsagar, D. B., & Waghchaure, S. V. (2014). Handwritten devanagari character recognition using neural network. *IOSR Journal of Computer Engineering.*
7. Trier, O. D., Jain, A. K., & Taxt, T. (1996). Feature extraction methods for character recognition–a survey.
8. Almualim, H., & Yamaguchi, S. (1992). A method for recognition of cursive handwriting. *IEEE Transactions on Pattern and Machine Intelligence.*
9. Krevat, E., & Cuzzillo, E. (2006). Improving off-line handwritten character recognition with hidden markov models. *Transaction on Pattern Analysis and Machine Learning*.
10. Abuhaiba, I. S. I., & Mahmoud, S. A. Recognition of handwritten cursive arabic characters. *IEEE Transaction on PA & MI*.
11. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. Gradient-based learning applied to document recognition. *Intelligent Signal Processin*g.
12. Kato, N., Suzuki, M., & Omachi, S. (1999). A handwritten character recognition system using directional element feature and asymmetric mahalanobis distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*
13. Sutha, J., & Ramraj, N. (2007). Neural network based offline tamil handwritten character recognition system. In *IEEE International Conference on Computational Intelligence and Multimedia Application.*
14. Singh, D., Dutta, M., & Singh, S. H. (2009). Comparative analysis of handwritten hindi character recognition technique. In *IEEE International Advanced Computing Conference (IACC'09).*

A Novel Ontology Design and Comparative Analysis of Various Retrieval Schemes on Education Domain in Protégé

Neera Chaudhary, Suresh Kumar, and Sachi Gupta

Abstract The goal of Semantic Web ensures that the computer is able to recognize the Web data and thus can offer people with a variety of intelligence service. Linked data is the data of today's web. As time passes, the world is changing and the technologies are developing in the same way. The computers are developing, and from the isolated version, they have been entered in the field of networks of information exchange. As the data available on the Web is increasing, it is becoming difficult to handle it and retrieve it relevantly, so Semantic Web helps to search semantically as data on it is available in form of linked data (ontology).In this paper, we have proposed an ontology on education domain with the help of Protégé tool.

Keywords Semantic Web · Ontology · Query

1 Introduction

In the present environment, we can transform a massive library of interlinked documents, via computers, and present to people. Basically it has grown from hypertext system so everyone can contribute or share to it. The drawback is that liability and the quality of information cannot be same or guaranteed [1]. So, there is a need to extent the version of HTML. This was the main reason for defining another version called XML (extensible markup language) which has the task-specific extension and arbitrary domains. Semantic Web (SW) is an XML application, whose prime objective is to develop and make the present Web more semantically richer [1]. From the time of

N. Chaudhary (✉) · S. Gupta
Raj Kumar Goel Institute of Technology Ghaziabad, Uttar Pradesh, India
e-mail: neerachaudhar18@gmail.com

S. Gupta
e-mail: sachifcs@rkgit.edu.in

S. Kumar
Ambedkar Institute of Advanced Communication Technologies & Research Government of NCT of Delhi, New Delhi, India
e-mail: sureshpoonia@aiactr.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_48

the 1990s, AI researcher communities have investigated that ontology has become a thrust area of Semantic Web research [1] that helps to provide shared and common understanding of some specific domain and make deliberate communication between people and application system. This reason made ontology more popular [1].

Semantic Web consists of formal Resource Description Framework (RDF), and it helps to overcome the limitation of understanding information that is faced by the current web. It also consists of a set of languages known as OWL, and it is based on RDF models. Vocabularies are used in Semantic Web which describes and explains the concepts and relationship that are related to a specific area of concern. A vocabulary can be considered as a special form of ontology, sometimes also merely as a collection of URIs with a described meaning. Link data is current Web data. It is publicly available in RDF data, identified by URI. In linked data, data is linked to one another.

The entire idea of ontology may sound alike to the concept of RDF. Every ontology is an RDF graph [2]. Ontology can be defined in several ways. There are several definitions of ontology. One of the most popular definition is ontology as "An explicit specification of conceptualization" [2, 3]. The SW depends heavily on the formal ontologies that structure underlined data for the purpose of comprehensive and transportable machine understanding. Ontology is the branch of philosophy that studies the nature of existence and the structure of reality [4]. The term ontology is the backbone of Semantic Web. It gives a concise and systematic means for defining the semantics of Web resources and describes the relevant domain concepts and properties of those concepts [2]. The Semantic Web understands the content of Web resources and combines and relates the content of other resources so we need a system which should be able to interpret the semantics of each resource so that it can represent the content of those resource [3, 4].

2 Designing Proposed Ontology

Education domain is one of the major areas of research today. It consists of a group of institutions such as schools, colleges, universities, ministries of education and teacher training institutions. Its prime objective is to provide right to education to all children as well as young people. It covers a wide range of people, for example, students, teachers of school and professors of university, colleges, etc. [4, 5].

The main objective to design this ontology is to integrate society which interacts a large number of stakeholders like schools, colleges and universities. Moreover, it can be extended to include parents, local communities like MCD schools, RWA institutions and pre-nursery and primary schools [6].

Designing an ontology depends on various components such as class, sub-class, individuals, attributes, relations, functions and restriction terms, rules, axioms and events. The common ontology engineering methodology [7] containing several steps is mentioned below which helps to design and share structural similarities of an effective ontology of any domain.

2.1 Ontology Engineering Process

Ontology engineering process basically contains the following five steps in ontology development [7].

Step-1: Selection of domain.
Step-2: Define classes and sub-class related to the domain.
Step-3: Define properties (Object/Data) of each class and sub-class as much possible.
Step-4: Define Individual or instance of each of the objects.
Step-5: Make relationship among classes and sub-classes.

2.2 Ontology for Education Domain

All the above steps are linked to each other. Protégé tool provides a flexibility to design an ontology. Follow the above steps. We design an ontology for education domain. The proposed ontology is mentioned in the following Fig. 1.

The designed ontology is saved in OWL extension file format. For this, it uses two main Semantic Web standards RDF and OWL. Resource Description Framework (RDF) specifies the metadata and describes Web resources which is processed by machines. RDF has many applications in search engines, pattern recognition and retrievals. There are three parameters in RDF, i.e., subject, predicate and object called triplets. OWL is very rich with respect to properties. It is the extension of RDF and has additional properties also to design ontologies in more effective manner [1, 2].

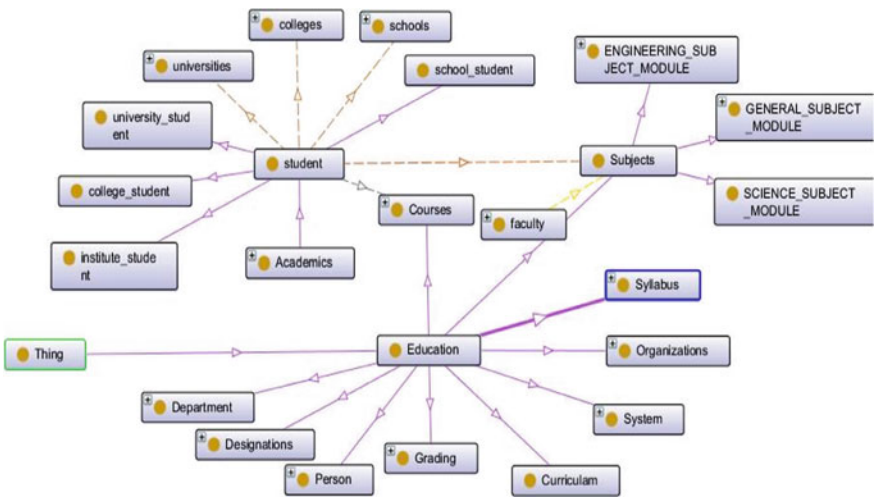


Fig. 1 Proposed ontology for education domain

3 Ontology Analysis

The proposed ontology has number of objects as well as data properties used with its domain and ranges. The following Table 1 represents the limited version of property, domain and ranges which is mentioned below.

Object properties map instance to other instance. Data properties are related to instance to literal data/constant value. In the case of the domain property, when we map a subject to an object using a property with the associated attribute, then the subject qualifies as a type of thing which specifies in the domain. The range works exactly like the domain, but it is applied to the object of the statement and not the subject [8, 9].

After designing the proposed ontology, there is an ontology matrices between the % of count of each matrices components and matrices are mentioned in the following Fig. 2. The figure shows that in our proposed ontology, class axiom is used almost 50.17%, Logical Axioms—25.41%, Class Count—15.84%, Object Property—1.82%, Data Property—1.82% and Individuals—4.95% in education domain.

Table 1 Property-domain-range description

Object property			Data property
Property	Domain	Range	
admitted_in	Student	School, university, college	topDataProperty course empno fcontact fdepartment fdesignation fname saddreess scontact sname ssn ssubject
affiliated_to	Student	School, university, college	
has	Faculty	Department	
has_department	Student, faculty	Department	
has_designation	Faculty	School, university, college	
has_faculty	Department	School, university, college	
opt	Faculty	Subject, course	
regiters_in	Student	Course	
studies	Student	Subjects	
teaches	Faculty	Courses, subjects	

Fig. 2 Ontology matrices

4 Empirical Results

Query language used for making queries in search engine search index is used to retrieve information. Formally Query language is defined in a context-free grammar (CFG). Query language tools [10] are used by users in textual, visual/UI or speech form. Query language mostly used forms to search are keyword and sentence based. In this paper, we have used keyword search mechanism.

To find the relevancy of retrieved result in protégé, we have done the analysis of three (03) Query languages which are OntoGraf Query, DL Query and SPARQL Query. These languages are like in-built tool in protégé that empower protégé users to Query the ontology data set and retrieve information. With the help of these Query tools, precise queries can be constructed using different searching criteria (searching keywords). In this paper, the table of all the three Query languages (QL) retrieved result using the same five (05) search keywords that are Student, faculty, Course, Subject, Department and Organization, in correspondence to the criteria used for searching by each QL is explained.

4.1 OntoGraf Query Analysis

OntoGraf [11] supports OWL ontologies for interactively navigating relationships. Table 2 describes the OntoGraf Query result description. Table 2 represents five search term or keywords used for searching. OntoGraf Query search is done based on the following five criteria mentioned below:

- Contains—In this, information containing the search terms are retrieved.
- Start With—In this, information starting with the search terms are retrieved.
- Ends With—In this, information ending with the search terms are retrieved.
- Exact Match—In this, information exactly matching with the search term are retrieved.

Table 2 OntoGraf Query result description

Search term	Criteria for OntoGraf Query search					Result
	Contains	Start with	Ends with	Exact match	Regular expression	
Student	8	4	5	1	8	26
Faculty	7	3	5	1	7	23
Course	7	3	4	0	7	21
Subject	4	1	0	0	4	09
Department	5	1	5	1	5	17
Organization	1	1	1	1	1	05

Table 3 DL Query result description

Search term	Criteria for DL Query search						Result
	Super classes	Ancestor classes	Equivalent classes	Sub-classes	Descendant classes	Individuals	
Student	1	4	1	4	4	2	16
Faculty	1	4	1	4	4	1	15
Course	1	2	1	4	4	1	13
Subject	1	2	1	3	10	5	22
Department	1	2	1	4	4	0	12
Organization	1	2	1	8	31	3	46

- Regular Expression—In this, information in the form of sequence of characters matching with the search terms are retrieved.

Based on the above explained criteria, a number of result retrieved from the novel designed ontology on education domain are mentioned in Table 2.

4.2 DL Query Analysis

DL Query [12] in protégé is used for searching a classified ontology. Table 3 describes the DL Query result description. Table 3 also represents the same five search terms or keywords used for searching. DL Query search is done based on the following six criteria explained and described below:

- Super classes—Retrieves the super class related to search term.
- Ancestor classes—Retrieves all the super classes related to search term.
- Equivalent classes—All the classes equivalent to search term are retrieved.
- Sub-classes—Sub-class that directly relates to the search term is retrieved.
- Descendant classes—All the sub-classes that relates to the search term are retrieved.
- Individuals—Individuals related to class belonging to search term are retrieved.

Based on the above explained criteria, a number of result retrieved from the novel designed ontology on education domain are mentioned in Table 3.

4.3 SPARQL Analysis

SPARQL [13, 14] is the Query language of the Semantic Web. It is an in-built tool for querying from ontology. It retrieves the user Query result in triple—subject, predicate and object. Based on the proposed education ontology, we apply command

Fig. 3 Query result in response to the user Query

Table 4 SPARQL Query result description

Search Term	Subjects	Objects	Results
Student	4	1	5
Faculty	4	1	5
Course	4	1	5
Subject	3	1	4
Department	4	1	5
Organization	6	1	7

and generated the Query results. The sample of Query in SPARQL is mentioned in the following Fig. 3 which represents the Query result output.

The following Table 3 mentioned the result based on the search terms in SPARQL. Here, based on the search terms, subjects' frequencies are presented by individual objects. The results are retrieved on the bases of one to one, one to many and many to many relationships (Table 4).

4.4 Overall Analysis

Various users are biased to use SPARQL, DL Query, OntoGraf Query on the basis of results extracted by these tools for knowing their education information. It was observed from different empirical result tables that SPARQL has less retrieved results in comparison to other Query tools like OntoGraf and DL Query. The fact that SPARQL could be far appropriately suited in education-based ontology for extracting imperative information rather than the other knowledge builder tools could be concluded by analyzing the analogy in Fig. 4.

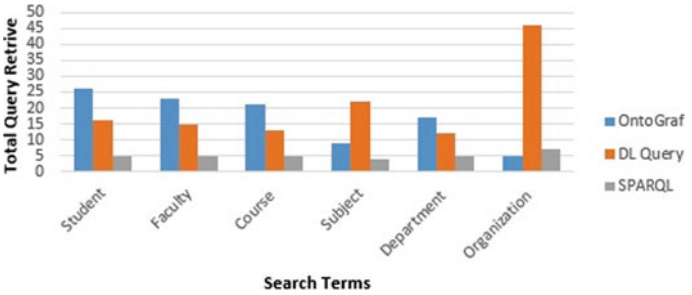


Fig. 4 Overall analysis of OntoGraf, DL and SPARQL Query result

5 Conclusion and Future Scope

In this paper, we have proposed an ontology on education domain. We have performed six queries (keywords) [15] on this ontology using three tools. Then after, we have compared the results of these three tools. In our empirical analysis, we have concluded that the results obtained through SPARQL tool are better than DL Query and OntoGraf Query tool. As the result retrieve through DL Query and OntoGraf Query tool provides semantically rich response but these are not much expressive whereas SPARQL tool provides semantically as well as more expressive power of producing the result.

For future, we have planned to use results of all these tools as an application for ranking. We are developing an effective mathematical method for ranking in information retrieval, which will be published in near future.

References

1. Giri, K. (2011). Role of ontology in semantic web. *DESIDOC Journal of Liberty & Information Technology, 31*(2), 116–120.
2. Synak, M., Dabrowski, M., & Kruk, S. R. Semantic Web and Ontologies. In *Semantic digital libraries,* (pp. 41–54).
3. Bamashmoos, F., Holyer, I., Tryfonas, T., & Woznowski, P. (2017). Towards secure SPARQL queries in semantic web applications using PHP. In *2017 IEEE 11th International Conference on Semantic Computing.*
4. Anand, S., & Verrma, A. (2010). Development of ontology for smart hospital and implementation using UML and RDF. *IJCSI International Journal of Computer Science Issues, 7*(5), 206.
5. Ceron-Figueroa, S., Lopez-Yanez, I., Alhalabi, W., Camacho-Nieto, O., Villuendas-Rey, Y., Aldape-Perez, M., et al. (2017). Instance-Based ontology matching for e -learning material using an associative pattern classifier. *Computers in Human Behavior, 69,* 218–225. https://doi.org/10.1016/j.chb.2016.12.039.
6. Jacob, E. K. Ontologies and the semantic web. In *April/May 2003-Bulletin of the American Society for Information Science and Technology.*

7. Liu, Y., Wang, Y., & Yang, D. (2014). DLOSS: An improved ontology-based semantic web search system. In *Fifth International Conference on Intelligent Control and Information Processing,* (pp. 18–20). China: Dalian, Liaoning.
8. http://protegeproject.github.io/protege/views/object-property-characteristics/.
9. https://www.w3.org/TR/owl-ref/.
10. Goncalves, R. S., Tu, S. W., Nyulas, C. I., Tierney, M. J., & Musen, M. A. (2017). An ontology-driven tool for structured data acquisition using Web forms, Goncalves et al. *Journal Biomedical Semantics, 8,* 26 https://doi.org/10.1186/s13326-017-0133-1.
11. https://protegewiki.stanford.edu/wiki/OntoGraf.
12. https://protegewiki.stanford.edu/wiki/DLQueryTab.
13. https://www.w3.org/TR/rdf-sparql-query/.
14. https://en.wikipedia.org/wiki/SPARQL.
15. Singh, P. K., Agarwal, G., & Gupta, S. (2013). A new ranking technique for ranking phase of search engine: Size based ranking algorithm (SBRA). *International Journal of Computer Applications, 82*(5), 0975–8887.

Similarity Analysis of Legal Documents: A Survey

Anya Trivedi, Aashka Trivedi, Sourabh Varshney, Vidhey Joshipura, Rupa Mehta, and Jenish Dhanani

Abstract A non-trivial task in the legal domain is grouping together similar legal case documents. This becomes particularly necessary in judiciaries following a Common Law System, as the judgements of previous similar cases prominently affect the outcome of a present case. A common approach of analysing the similarity of documents is by considering the documents as a whole. But, because legal case documents are extremely verbose, it may be beneficial to analyse the similarities of only the important concepts, or the summary, of the documents at hand. The primary focus of this survey is to explore the different methods used to analyse the similarity between legal case documents. This paper categorizes these efforts into citation-based, content-based or summary-based methods. A thorough survey of the existing methods of summarizing legal documents has also been explored.

Keywords Legal document · Document similarity · Text summarization · Text classification

A. Trivedi (✉) · A. Trivedi · S. Varshney · V. Joshipura · R. Mehta · J. Dhanani
S.V. National Institute of Technology, Surat, Gujarat 395007, India
e-mail: trivedianya@gmail.com

A. Trivedi
e-mail: aashka.trivedi@gmail.com

S. Varshney
e-mail: sourabvarshney111@gmail.com

V. Joshipura
e-mail: vidheyjoshipura98@gmail.com

R. Mehta
e-mail: rgm@coed.svnit.ac.in

J. Dhanani
e-mail: jenishdhanani26@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_49

1 Introduction

The legal domain is often concerned with retrieving highly diverse information from a large document corpus. Document retrieval and analysis tasks are extremely important to practitioners of the *Common Law Judicial System* [22]. Here, one judgement of a case sets precedence over all future judgements. The challenges that arise with working with a legal document corpus is not only due to its enormous size, but the nature of the legal document itself. Legal case documents are highly verbose and are often written with many redundant words [14, 16]. Moreover, many legal documents do not contain an explicitly labelled summary. Thus, an automated document summarization system would be useful in identifying the important concepts of a legal document. This summary can also be used in effectively analysing other legal documents similar to the one at hand.

This review discusses the existing works in the field of document similarity analysis and automated text summarization, with a focus on its applications in the legal domain. This survey classifies these efforts into three broad approaches, as shown in Fig. 1. In **citation-based methods**, similar documents are considered to contain a minimum threshold number of common in-citations or out-citations. This type of similarity analysis works on the intuition that similar documents should cite, or be cited by, the same set of cases. **Content-based methods** analyse the entire document's content to gauge similarity. This is done either by comparing the similarity between the vectorized version of the text, or the inherent topics present within it. A summary of a given document expresses all the inherent topics present within it in a shorter manner. In **summary-based methods**, a legal document's summary can be either generated using automatic summarization methods, or by identifying the paragraphs that act as a summary of the entire document. The same text-based or topic-based approaches used for entire documents can then be applied onto these summaries.

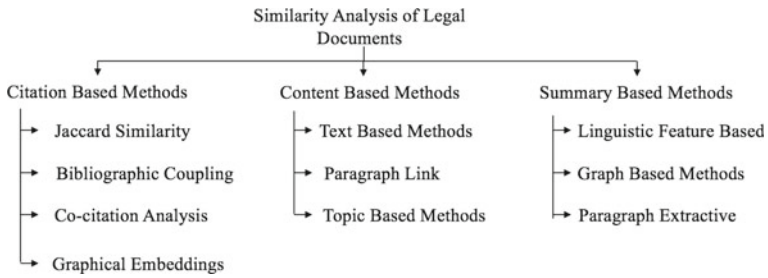


Fig. 1 Similarity analysis of legal documents

2 Citation-Based Similarity Measures

Network-based similarity measures, as proposed by Kumar et al. [17], comprise of creating a network of legal documents connected by their citations. Citations can be classified as either in-citations or out-citations depending on the relationship of the citation with the given document. The **in-citations** for a given judgement are those case citations, present in other judgements, which are referring to the given judgement. The **out-citations** of a given judgement are the case citations (of other judgements) that the present judgement refers to while passing the court's decision [17].

A network of legal documents is represented as a directed graph such that nodes represent the documents (where the set V of vertices is our corpus). A directed edge from vertex D1 to vertex D2 implies that the document D1 cites the document D2 [4, 21]. This relationship can be seen in Fig. 2. Here, D1, D2, D3, D4 and D5 are legal documents, and a directed edge between nodes denotes their citations. The out-citations of D1 are D2, D3 and D4, while the in-citations of D4 are D1, D2 and D5. Two documents can simultaneously linked directly (D1-D4) and indirectly (D1-D2-D4). A citation graph is helpful in determining the similarity of case documents and can also be used to explore the reachability of a case [26, 28].

Citation-based similarity methods consist of Jaccard similarity, bibliographic coupling, co-citation and graphical embeddings.

2.1 Jaccard Similarity

Wagh and Anand [28] create a weighted graph, where an edge between two document nodes exists if they are similar. The weight of the edge is obtained by calculating the Jaccard similarity between the set of case-citations(out-citations) of the two cases. This is done by calculating the ratio of their intersection to their union [21, 28].

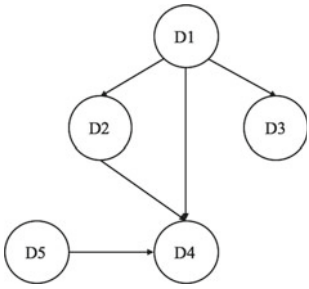


Fig. 2 Network of legal documents

2.2 Bibliographic Coupling

Bibliographic coupling is a link-based similarity measure that is concerned with the out-citations of a document. Two documents are said to be bibliographically coupled if they refer to one or more common works [16]. The two documents are determined to be similar if their bibliographic coupling is greater than a threshold value, δ. Given two documents D1 and D2, and their set of out-citations OC1 and OC2, respectively, the bibliographic coupling between them is defined as the number of common out-citations, or:

$$\mathrm{B}(\mathrm{D1}, \mathrm{D2}) = \mathrm{OC1} \cap \mathrm{OC2} \tag{1}$$

2.3 Co-citation Analysis

While bibliographic coupling links out-citations, co-citation links the number of documents which cite the two given documents together, i.e. their in-citations. Thus, if a document *D* cites two documents D1 andD2 together, D1 and D2 are said to be co-cited. The more co-cited the two documents are, the higher the likelihood that they are semantically similar [26]. The two documents are determined as similar if their co-citation is greater than a threshold value, δ. Given two documents D1 and D2, and their set of in-citations IC1 and IC2, respectively, the co-citations between them are the number of common in-citations, or:

$$\mathrm{C}(\mathrm{D1}, \mathrm{D2}) = \mathrm{IC1} \cap \mathrm{IC2} \tag{2}$$

2.4 Graphical Embeddings

Bhattacharya et. al. [4] calculate the "Precedent Citation Similarity" through the use of Node2Vec. Node2Vec [12] maps the nodes of a graph to a vector space such that the network neighbourhoods of the nodes are preserved. Thus, nodes with similar neighbours will have similar vector embeddings.

A major factor affecting citation-based similarity measures is that citation graphs for legal documents tend to be sparse [4]. Hence, co-citation methods may not be as accurate as graph embedding-based methods [4] or bibliographic coupling [16, 17]. An advantage provided by citation-based similarity approaches is that the network made for the same can also be used to explore the reachability of one case from the other. Wagh and Anand [28] show that this may further be used to determine whether cases are equivalent or not.

3 Content-Based Similarity

Documents can be viewed as vectors of text or a collection of topics. Thus, by viewing the document as a whole, we discuss the following three approaches to determining similarity.

3.1 Text-Based Measures

Here the content of a document is represented as a vector in vector space [19]. Similar documents will have similar embeddings, i.e, vectors that will be closer in the vector space. Each component of the document vector corresponds to a unique term from the vocabulary. It may also contain the TF-IDF score [2] for the corresponding term. After acquiring the vectors of two documents, their corresponding cosine similarities are calculated to determine their similarity [21]. A cosine similarity closer to 1 denotes a smaller angle between the vectors and hence higher similarity. Kumar et al. [17] have explored two ways of defining the scope of the vocabulary for the analysis of legal documents. **All terms** vocabulary is defined as the set of all the distinct terms occurring in the document corpus [17]. **Legal terms** vocabulary is the set that contains those distinct terms that occur in the document corpus, as well as in a legal dictionary [17].

3.2 Paragraph Link

As observed by Kumar [16], link-based similarity measures are more effective for finding legal judgement than text-based similarity measures. However, links exist only sparsely in legal judgements and hence are not able to fetch more than a few more judgements which are similar to the given judgement. In order to efficiently use link-based methods, while overcoming the problem of a sparse citation graph, Kumar [16] proposes the use of a "paragraph link". A paragraph link between two paragraphs of different documents is said to exist if they are found to be similar by some similarity measure. Bibliographic coupling is then performed on the two documents, comparing paragraph links instead of out-citations. If the number of common paragraph links between the two documents above a threshold, the documents are considered to be similar. This showed the highest correspondence with legal expert opinions regarding similar legal documents.

3.3 Topic-Based Similarity Measures

Topic modelling methods use statistical methods to determine the underlying topics of a textual document. Mandal et al. [21] and Kumar and Raghuveer [27] proposed to use topic modelling to find similar documents, by using Latent Dirichlet allocation [6] to obtain a fixed length vector for each document (where the length corresponds to number of topics inherently present in the corpus). The resultant vector reflects the distribution of topics in the text [6], and the cosine similarity of these vectors are calculated to determine which documents are similar based on the topics that they contain. Building on the hypothesis of Sarvanan et al. [23, 24], Bhattacharya et al. [5] proposes the idea of thematic similarity. They use machine learning methods to identify 7 themes (or segments) of a legal document. In [4], segment level similarity for all segments of the required documents is calculated, and the final similarity score is given by computing either the average of all segment level scores or the maximum of them.

On a text-based level, it was found that comparing cosine similarity values based on legal terms was found to be more efficient than an all term cosine similarity analysis [17]. Moreover, [4, 21] find that embedding the entire document works well. The fact that the maximum aggregation worked better than the average aggregation strengthens the intuition that documents can be very similar with respect to certain themes (e.g. the facts), while differing with respect to other themes. This analysis further strengthens the idea that it may be fruitful to measure the similarity of only certain parts of the legal document, as opposed to the document as a whole.

4 Summarization of Legal Documents

Automatic summarization creates a concise summary of the document, which preserves all the necessary and important information of the original text [1, 3]. Text-based similarity measures at the term-level [2, 13, 17, 21], the paragraph-level [16], and topic-based similarity measures [6, 21] can then be applied to these generated summaries as opposed to the document as a whole.

There are two broad approaches to textual summarization: **Extractive Summarization** methods identify the most important sentences of the given text and extract them as they are. The sentences are picked up without any regard to domain knowledge of any kind [8]. **Abstractive Summarization** methods represent the important concepts in the given text in a way that differs from the original text, using some amount of domain knowledge to determine the importance of the various concepts [1, 3]. Because of the difficulties arising in natural language generation techniques, inference and the semantic nuances of analysing and generating text, extractive summarization methods often outperform abstractive methods [1].

Kanapala et al. [14] observed that the task of summarizing documents from the legal domain is different from that of other domains (such as scientific papers, news

articles, etc). This is because of their size, intensive vocabulary, ambiguity in text, differing internal structures and their extensive citations.

This paper classifies work done in the field of legal document summarization into linguistic feature-based approaches [9–11, 15, 24, 29], graph-based approaches [14, 23, 24], and paragraph extracting approaches [16, 25]. These have been discussed below.

4.1 Linguistic Feature-Based Summarization

Basic summarization methods use the information present at the sentence, document and collection (or citation) level to identify the major concepts present in the document. As observed by Galgani et al. [10] and Kanapala et al. [14], this alone may not be the case for summarizing legal documents, as the concepts of the document is also highly dependent on the context of the legal case. Galgani et al. [10] performed extractive summarization by proposing the creation of a knowledge base that specifies how these concepts should be used in different contexts. They identify important legal points from the case document, or "catch-phrases", which were identified through the use of Ripple Down Rules [7]. Galgani and Hoffmann [11] showed that this method for the analysis of legal citations outperforms the use of machine learning techniques for the same.

Wang et al. [29] propose the use of latent semantic analysis to select the best sentence description and term description of the topic. Since there are a limited number of topics in a legal text, Kanapala et al. [14] reason that an LDA-based approach would be more effective than an LSA-based approach in the summarization of legal documents. Kumar and Raghuveer [15] propose such a topic-based document summarization model, using latent Dirichlet allocation [6]. This work consists of an breaking the legal document into the seven topics as defined by Saravanan et al. [24]. Extractive summarization was performed by selecting the top two highest-scoring sentences for each topic. The final summary was presented by ordering the sentences according to their original order, as present in the document.

Galgani et al. [9] use in-citations and out-citations to generate catch-phrases from the legal document to prepare the summary. By performing a ROUGE-1 [20]-based evaluation, [9] showed that a citation-based approach performed better than an only text based one, without referencing any other legal cases.

4.2 Graph-Based Summarization

Graphical methods for the summarization of legal documents use sentences or paragraphs to represent the nodes of a graph. These nodes are ranked according to importance and similarity between the sentences or paragraphs and are then used to construct a summarization of the document [14]. Probabilistic models were applied to

automatically summarize documents related to the legal domain by Sarvanan et al. [24]. References [23, 24] used conditional random fields [18] to extract sentences through a term distribution model for summarizing a document. As observed by Kanapala et al. [14], the lack of freely available legal-text dictionaries prevents the semantic or sentiment study of the document pertaining to the legal domain. Such a study would allow the interconnection of words, sentences or paragraphs by determining their sentimental and linguistic polarities. Thus, summarization using only a graph-based approach may be unfruitful.

4.3 Paragraph Extractive Summarization

Schilder and Molina-Salgado [25] propose that in the case of legal documents, there exist paragraphs that act as a summary of the document. In order to find these "summary paragraphs", an inter-paragraph similarity score is created for every pair of paragraphs, based on the overall phrase similarity between the two. The paragraphs which are found to be the most similar to many other paragraphs are added to the extractive summary. This hypothesis is further substantiated by Kumar [16], who states that in the case of Indian Legal Documents, the headnote acts as the effective summary of the judgement. Hence, it may be fruitful to classify the paragraphs as headnote or other and use the classified headnote verbatim as the summary of the document.

5 Conclusion

This survey classifies work done for the similarity analysis of legal documents into three broad approaches. These are similarity on the basis of citations, similarity between legal documents taken as a whole, and similarity between the summaries of the legal documents. A broad observation made is that most prior works agree that using parts of the document (either citations, specific paragraphs, or a summary) for similarity analysis may be more efficient than considering the document as a whole. This may because legal documents often contain redundancies. Additionally, the use of certain phrases may be mandated for legal documents of a specific court, thus skewing similarity results. Moreover, certain words may be reflective of the author's writing style and thus may be present in all the author's works, irrelevant of whether the case is similar or not.

Legal citations can be used to understand the relationship between two cases. The intuition behind this is that similar legal documents would cite the same cases. Wagh and Anand [28] also showed that a citation-based similarity approach using Jaccard similarity was closer to human opinion and provided more insights into the similarity of two documents as compared to a text-based approach using cosine similarity. The major limitation of using citation-based methods is that the citation graph of legal

documents tends to be sparse [17]. Bhattacharya et al. [4] also explored a hybrid approach to improve similarity analysis by combining textual similarity methods with citation-based similarity methods to yield better results. As proposed by Kumar [16], the use of paragraph links in bibliographic coupling helps to overcome the sparseness of citation graphs that negatively affect pure citation-based methods, while taking advantage of the performance of link-based similarity methods. This idea of using only specific paragraphs to gauge similarity is very similar to the work of Bhattacharya et. al [5].

References [4, 21] found that documents can be similar with respect to certain topics or themes, further strengthening the intuition that similarity can be measured by considering only certain parts of the legal document. The properties of legal documents can be exploited to generate accurate extractive summaries. Linguistic feature-based summaries, while effective in capturing the important concepts of the document succinctly, might incur the overhead of building a knowledge base [10], or of identifying the topics of the document [15]. Graph-based summarization methods rank nodes according to importance and inter-sentence similarity. But the lack of accessible legal dictionaries may reduce the efficiency of such summarization methods [14]. Paragraph extractive summarization methods extract relevant paragraphs in order to prepare a summary. However, for legal documents spanning multiple paragraphs, computing inter-paragraph similarity in order to prepare a summary [25] may be computationally expensive.

References

1. Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). *Text summarization techniques: A brief survey*.
2. Bafna, P., Pramod, D., & Vaidya, A. (2016). Document clustering: Tf-idf approach. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)* (pp. 61–662016). https://doi.org/10.1109/ICEEOT.2016.7754750.
3. Bhatia, N., & Jaiswal, A. (2015) Literature review on automatic text summarization: Single and multiple summarizations. *International Journal of Computer Applications, 117*(6), 25–29 (full text available).
4. Bhattacharya, P., Ghosh, K., Pal, A., & Ghosh, S. (2020). *Methods for computing legal document similarity: A comparative study*.
5. Bhattacharya, P., Paul, S., Ghosh, K., Ghosh, S., & Wyner, A. (2019). *Identification of rhetorical roles of sentences in Indian legal judgments*. arXiv preprint arXiv:1911.05405.
6. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022. http://dl.acm.org/citation.cfm?id=944919.944937.
7. Compton, P., & Jansen, R. (1988). Knowledge in context: A strategy for expert system maintenance. In *Australian Joint Conference on Artificial Intelligence* (pp. 292–306). Springer.
8. Cranganu-Cretu, B., Chen, Z., Uchimoto, T., & Miya, K. (2002). Automatic text summarizing based on sentence extraction: A statistical approach. *International Journal of Applied Electromagnetics and Mechanics*, *13*(1–4), 19–23.
9. Galgani, F., Compton, P., & Hoffmann, A. (2013). Citation based summarisation of legal texts. In *Pacific Rim International Conference on Artificial Intelligence* (pp. 40–52). Springer.

10. Galgani, F., Compton, P., & Hoffmann, A. (2012). Combining different summarization techniques for legal text. In *Proceedings of the Workshop on Innovative Hybrid Approaches to the Processing of Textual Data* (pp. 115–123). Association for Computational Linguistics.
11. Galgani, F., & Hoffmann, A. (2010). Lexa: Towards automatic legal citation classification. In *Australasian Joint Conference on Artificial Intelligence* (pp. 445–454). Springer.
12. Grover, A., & Leskovec, J. (2016). *node2vec: Scalable feature learning for networks*.
13. Huang, A. (2008). *Similarity measures for text document clustering*.
14. Kanapala, A., Pal, S., & Pamula, R. (2017). Text summarization from legal documents: A survey. *Artificial Intelligence Review*, *51*, 371–402.
15. Kumar, R., & Raghuveer, K. (2012). Legal document summarization using latent dirichlet allocation. *International Journal of Computer Science and Telecommunications*, *3*, 114–117.
16. Kumar, S. (2014). *Similarity analysis of legal judgments and applying 'Paragraph-link' to find similar legal judgments*. Ph.D. thesis, International Institute of Information Technology Hyderabad.
17. Kumar, S., Krishna Reddy, P., Reddy, V., & Singh, A. (2011). Similarity analysis of legal judgments. In *Proceedings of the 4th Annual ACM Bangalore Conference* (p. 17). https://doi.org/10.1145/1980422.1980439.
18. Lafferty, J., Mccallum, A., & Pereira, F. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning* (pp. 282–289).
19. Le, Q. V., Mikolov, T. (2014). *Distributed representations of sentences and documents*.
20. Lin, C. Y. (2004). Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out* (pp. 74–81).
21. Mandal, A., Chaki, R., Saha, S., Ghosh, K., Pal, A., & Ghosh, S. (2017). Measuring similarity among legal court case documents. In *Proceedings of the 10th Annual ACM India Compute Conference* (pp. 1–9). Compute '17, ACM, New York, NY, USA. https://doi.org/10.1145/3140107.3140119, http://doi.acm.org/10.1145/3140107.3140119.
22. Pejovic, C. (2001). Civil law and common law: Two different paths leading to the same goal. *Poredbeno Pomorsko Pravo, 40*, 7–32. https://doi.org/10.26686/vuwlr.v32i3.5873.
23. Saravanan, M., Ravindran, B., & Raman, S. (2008). Automatic identification of rhetorical roles using conditional random fields for legal document summarization. In *Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-I*. https://www.aclweb.org/anthology/I08-1063.
24. Saravanan, M., Ravindran, B., & Raman, S. (2006). Improving legal document summarization using graphical models. *Frontiers in Artificial Intelligence and Applications*, *152*, 51.
25. Schilder, F., & Molina-Salgado, H. (2006). Evaluating a summarizer for legal text with a large text collection. In *3rd Midwestern Computational Linguistics Colloquium (MCLC)*. Citeseer.
26. Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*(4), 265–269. https://doi.org/10.1002/asi.4630240406, https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.4630240406.
27. Raghuveer, K. (2012). Legal documents clustering using latent dirichlet allocation. *International Journal of Applied Information Systems, 2*(6), 27–33. Published by Foundation of Computer Science, New York, USA.
28. Wagh, R., & Anand, D. (2017). Application of citation network analysis for improved similarity index estimation of legal case documents: A study. In *2017 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC)* (pp. 1–5).
29. Wang, Y., & Ma, J. (2013). A comprehensive method for text summarization based on latent semantic analysis. In *Natural Language Processing and Chinese Computing* (pp. 394–401). Springer.

Performance Enhancement of STATCOM Integrated Wind Farm for Harmonics Mitigation Using Optimization Techniques

Mohamed Hamdy, Mahmoud A. Attia, Almoataz Y. Abdelaziz, Sachin Kumar, Kumari Sarita, and R. K. Saket

Abstract Renewable energy generation is growing every day all over the world. This injects the grids with harmonics and increases the total harmonics distortion of the systems. In the other hand, Flexible AC transmission systems (FACTS) are used in the different power systems for the enhancement of the stability of these systems. In this paper, Flexible AC transmission systems (FACTS) shall be used not for the enhancement of the stability of the system as usual, but to mitigate the harmonics of the system and decrease the Total harmonics distortion (THD). The Static Synchronous Compensator (STATCOM) performance is compared using The Harmony Search Optimization Algorithm (HSA) and the Invasive Weed Optimization (IWO) trying to achieve better results. MATLAB/SIMULINK is used to create a power system model of wind generation system and then is used to compare the two techniques.

Keywords Renewable energy · Wind turbines · Harmonics · Total harmonics distortion · Optimization techniques · Flexible AC transmission systems · Static compensator

1 Introduction

The harmonics generated from the wind turbine are one of the withdraws of the wind generation as it causes lots of problems in the power systems. The harmonics can be defined as a function that has integer frequencies are multiples of the frequency of

M. Hamdy · M. A. Attia
Department of Electrical Power and Machines, Faculty of Engineering, Ain Shams University, Cairo, Egypt

A. Y. Abdelaziz
Faculty of Engineering & Technology, Future University in Egypt, Cairo, Egypt
e-mail: almoatazabdelaziz@hotmail.com

S. Kumar (✉) · K. Sarita · R. K. Saket
Department of Electrical Engineering, IIT (BHU), Varanasi, Uttar Pradesh, India
e-mail: sachinkumar.rs.eee18@itbhu.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_50

the fundamental power waveform. Some of the electrical components of the power system such as the power panels and transformers can become resonant mechanically to the magnetic fields resulting from the harmonics with higher frequency when these frequencies of harmonics are prevalent. As a result of that, these components might make a buzzing noise and suffer from vibration corresponding to different frequencies of harmonics. In the modern power system, the harmonics from 3rd to the 25th can be considered as the most common range.

The total harmonic distortion (THD) is an important parameter that represents the harmonic distortion level of any voltage or current signals. So, taking voltage as an example, THD is an index to compare the harmonic voltage components with the fundamental element of the voltage signal, as per Eq. 1.

$$\mathrm{THD} = \frac{\sqrt{\sum_{h=2}^{n} V_h^2}}{V_1} = \frac{\sqrt{V_1^2 + V_2^2 + \cdots V_n^2}}{V_1} \tag{1}$$

where, the variable h is the number of harmonics of the signal, and n is the maximum harmonic order of voltage, $_1$ is the nominal system voltage at the fundamental frequency [1]. The (ANSI/IEEE 519-1992) standard lists are one of the most widely used guides for THD limits in any power grid system [2]. Thus, the critical objectives of any harmonic analysis are to assist the system, both design and installation in meeting (IEEE 519-1992) standards regarding limiting THD.

The Flexible AC Transmission Systems (FACTS) involves a group of power electronic devices which are developed for the applications including phase-shifting transformers, tap changers, reactive compensators, synchronous condensers, etc. [3]. The FACTS devices are used to control the parameters of the transmission line, such as line reactance and node voltages [4]. The use of semiconductor switches makes the FACTS devices much faster as compared to the conventional mechanical switches, but the FACTS technology is costly [5].

The dependency of the performance of industrial applications, including information technology and production engineering on the power quality has made this an important factor to achieve [6]. At the generator side, the power quality is defined as the ability of the generator to generate power at 50 Hz without any variations to that power, but in for the transmission and distribution level, it is defined as the ability of the voltage to stay within the limit of 5%. Another definition of the power quality is that it is the analysis, measure, and improvement of the voltage of the bus to maintain this voltage to be sinusoidal in a wave within the rated voltage and frequency [7].

The connectivity of renewable energy resources (RERs) with the transmission or distribution systems are increasing without any issue due to the availability of power electronic converters. These converters affect the system power quality and increase the harmonics at the common coupling point. The mitigation of these power quality issues is very important and challenging too [8, 9]. So, the use of RERs in the traditional electrical system resulted in changing the reliability, management, power quality, protection, and control policies of electrical utilizes. In other words, good power quality can be considered as a critical factor for a reliable power system.

However, the non-sinusoidal waveforms produced from different electronic devices and nonlinear loads will result in poor power quality systems [10, 11].

During the last few years, the FACTS technology has been used in power systems for power quality enhancement, voltage stabilization, power factor correction and harmonic mitigation [12, 13]. So, there is significant research that was trying to get the maximum energy that can be extracted from the RERs, and the results of implementing FACTS devices in the grids with RERs are encouraging [14, 15]. So, the implementation of different FACTS devices in any of the power systems shall have the effect of reducing different problems of power quality. The FACTS's basic concept is the real and reactive power flow control and also the voltage control in the power system by the usage of the high-voltage power electronics devices [16, 17]. The usage of the Static Synchronous Compensator (STATCOM) facilitates the connection of wind power to any power systems [18], helps in the improvement of the power system stability that contains wind energy sources [19], improve voltage ride-through, and regulate the power flow [20]. The Dynamic Distribution System Compensator (DDSC) enhances the stability of main buses of the system and the power quality of the whole system. It increases the capacity of the distribution feeder transmission [21]. As for the Power filter compensator (SFC), it improves the power factor, stabilizes the AC voltage, and enhances the power quality of the system [22]. The Static Synchronous Series Compensator (SSSC) helps in providing voltage ride-through and regulation of power flow [20]. The Static Var Compensator (SVC) improves the stability of the power system with PV energy source [23], the power quality of the system, and the transient stability [24]. For the Distribution-STATCOM, it improves the voltage profile of the power system that has distributed wind generation [25].

2 Optimization Techniques

Harmony Search Optimization Algorithm (HSA) can be considered as one of the meta-heuristic techniques that are used in the optimization of electrical power engineering problems [26]. This optimization algorithm is using random numerical simulations with certain algorithmic steps so that it can find engineering problems optimum solution. The HSA uses the concept of the musical performance that tries to reach a pleasing harmony for the audience to deal with the optimization problem and its objective function. The steps involved in HSA can be found in [27]. Invasive Weed Optimization (IWO) is another Optimization Algorithm that we can consider as an evolutionary meta-heuristic algorithm to solve different engineering optimization problems [28, 29]. To perform this algorithm, we should use the following steps [30]:

Population initializing: The finite number of populations that we can call them (seeds) shall be selected, and they all shall have random positions.

Spatial dispersal: These seeds that are produced newly shall be spread all over the search area randomly and grown to new plants. In this part, the algorithm randomness and adaptation that shall be provided, and this because of the fact the generated seeds shall be randomly distributed over the dimensional search space (d), and this distribution shall be by normally distributed random numbers that have mean equal to zero, but also, they have variable variance. While the simulation is working, an alteration that is nonlinear shall obtain performance that we can call satisfactory, which is shown in Eq. 2. That shows that.σ_{iter} is the present time step standard deviation, while $\text{iter}_{\max}$ is the maximum iterations number, and finally, n is the index of the nonlinear modulation.

$$\sigma_{\text{iter}} = \frac{(\text{iter}_{\max} - \text{iter})^n}{\text{iter}_{\max}^n}(\sigma_{\text{intial}} - \sigma_{\text{final}}) + \sigma_{\text{final}} \tag{2}$$

Competitive exclusion: In this process, the maximum plants' number shall be reached. After that, until reaching the maximum iterations number, the procedure shall continue, and this point, the plant that has the best fitness that is the closest to the optimal solution should be achieved hopefully.

Using the optimization technique to get the best (lowest) THD by controlling the control gains of the STATCOM by making them the parameters of the Optimization Technique, and making the objective function of the Optimization Technique, and choosing the algorithm to minimize the objective function (THD). So, the Optimization Technique shall use random values of the gains from the specified range that is inputted in it and gets the corresponding THD, and then changes the gains again and check the corresponding THD and compare it with old one. Finally, use the lower one, Optimization Technique shall change the gain back, and so on until we get the best (lowest) THD. The specified range of the gain shall be selected in the beginning to be from 0 to maximum by trial and error until reaching to the final range of gain, and then apply the Optimization Technique with this range, to get the best gains that can achieve the best (lowest) THD, Fig. 1 represents this process.

3 Results and Discussion

The model that is constructed by MATLAB SIMULINK and it can be divided to a wind farm that uses Induction Generators (IG), and that is 9 MW. The model consists of a wind farm that has six 1.5 MW wind turbines, and this wind farm is connected to the distribution system that is 25 kV, and using a 25 km, 25 kV feeder, this distribution system is connected to a 120 kV grid. Three pairs of 1.5 MW wind turbines simulate this 9 MW wind farm. The stator winding of the induction generators (IG) is directly connected to the 50 Hz grid, and the variable-pitch wind turbine is driving the rotor. To maintain the output power of the generator at the nominal value while winds speed

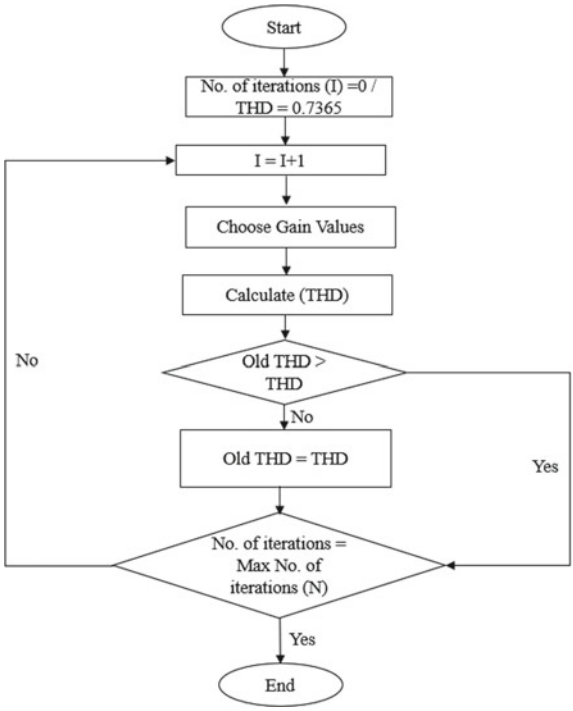


Fig. 1 The flowchart of THD calculation using optimization algorithm

exceeds the nominal value that is (9 m/s). The speed varies approximately between 1 per unit (PU) at no load and 1.005 per unit (PU) at full capacity.

The model is then modified with adding of 3 MVAr STATCOM to limit the total harmonics distortion (THD) of the system. The STATCOM is phasor type as the wind farm turbine so that it can work in the system. For the wind farm system without any harmonics reduction device, the THD is 73.65% which is very high and can cause lots of instability in the system. Figures 2 and 3 represent the wind farm power wave farm (MW) and the voltage waveform (pu), and we shall consider these waveforms as the standard to compare the other results with.

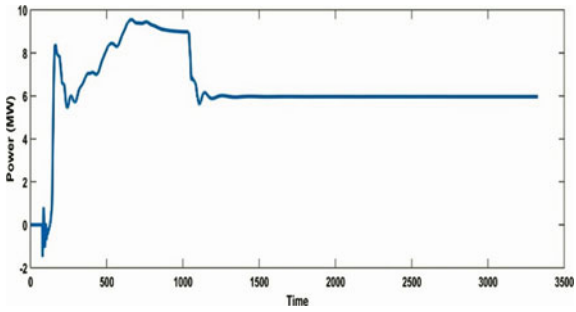


Fig. 2 Wind farm power (MW)

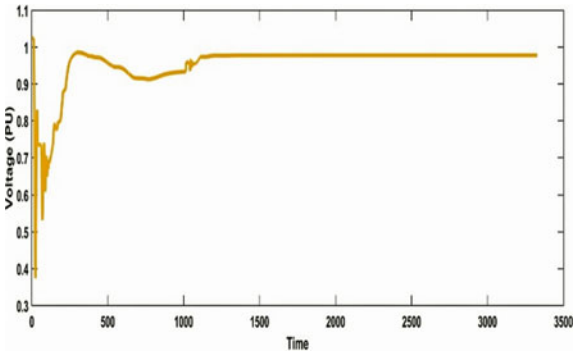


Fig. 3 Wind farm (abc) voltage (PU)

The wind farm system that is equipped with STATCOM that uses the HSA has much better THD as the (THD) is 0.7734% which is a big improvement from the 73.65% (THD) of the normal system, which is 98.95% Improvement from the normal system as shown in Fig. 4a. The power waveform (MW) of the wind farm system that is equipped with STATCOM, it becomes more stable, and Fig. 5 represents the waveform of it. Figure 6 shows the voltage waveform (PU) of the system, which is far more stable than the voltage of the normal system. The same model of the wind farm connected to the STATCOM is used again but with using IWO this time. The

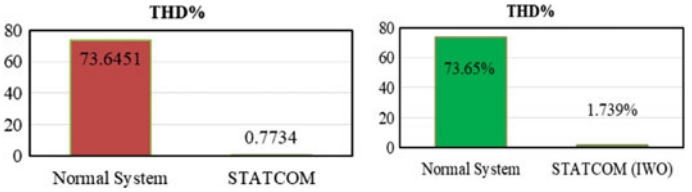


Fig. 4 Comparison between normal system and STATCOM with **a** HSA system THD, **b** IWO THD

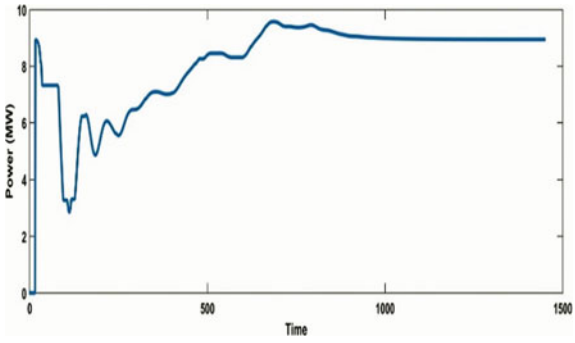


Fig. 5 Wind farm with STATCOM with HSA power (MW)

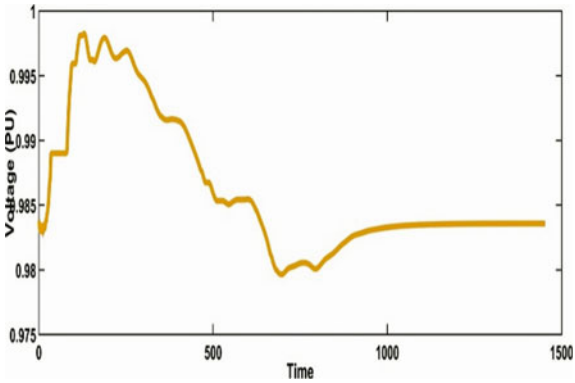


Fig. 6 Wind farm with STATCOM with HSA voltage (PU)

THD of the system became 1.739% by IWO which can also be considered as a big improvement from the 73.65% (THD) of the normal system, as it is 97.25% Improvement from the normal system as shown in Fig. 4b. The power waveform (MW), it becomes slightly more stable, and Fig. 7 represents the waveform of it. Figure 8 shows the voltage waveform (PU) of the system, which is also slight stable than the

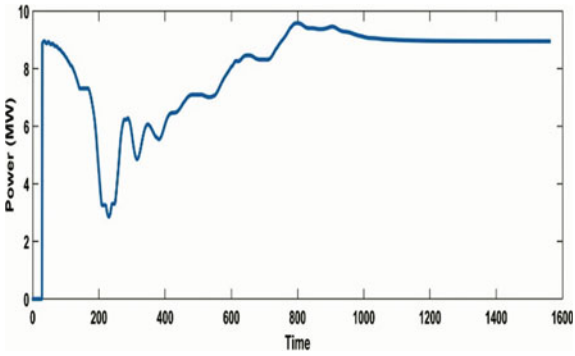


Fig. 7 Wind farm with STATCOM using IWO power (MW)

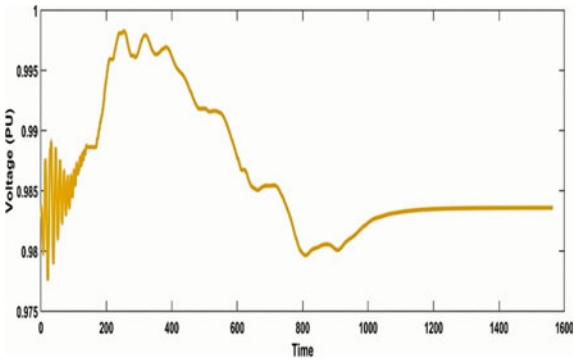


Fig. 8 Wind farm with STATCOM using IWO voltage (PU)

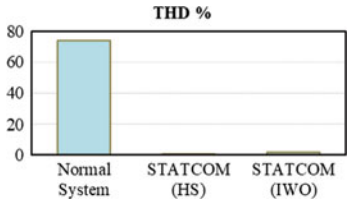


Fig. 9 Comparison between normal system and STATCOM system with IWO and HSA techniques

voltage of the normal system. After testing both Techniques, the Harmony Search Algorithm (HSA) has better results than the IWO in the Total Harmonics Distortion (THD) improvement that is 98.95% for the Harmony Search Algorithm (HSA) while it is 97.25% for the IWO. The comparison regarding the Total Harmonics Distortion (THD) improvement is shown in Fig. 9. Figures 10 and 11 shows the difference between the effect of the STATCOM of both power and voltage waveforms using both Techniques, Harmony Search Algorithm (HSA) and IWO.

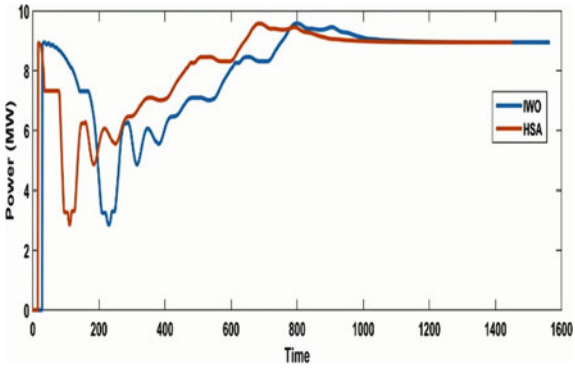


Fig. 10 Comparison between the system power using both IWO and HSA techniques

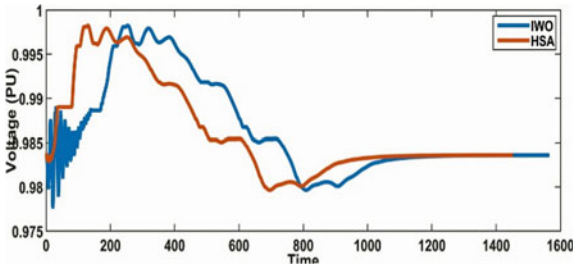


Fig. 11 Comparison between the system voltage using both IWO and HSA techniques

4 Conclusion

The paper has described the effect of STATCOM in the wind integrated power system. The output power, voltage and THD are analyzed in the wind farm model with STATCOM using HSA and IWO techniques. Both the HSA and the IWO are used to optimize the STATCOM performance in mitigation of the THDs in the power system integrated with a wind farm. The HSA has slightly better results than the IWO regarding the THD reduction also in the enhancement of voltage and power performance. The work can further be extended to observe the performance of the power system implementing other FACTS devices. Also, the use of other meta-heuristic optimization techniques can further be applied for better results.

References

1. Schwanz, D., & Leborgne, R. C. (2014). Comparative harmonic study of a wind farm: Time vs frequency domain simulation. In *International Conference on Harmonics and Quality of Power*. Bucharest, Romania: IEEE. https://doi.org/10.1109/ichqp.2014.6842774.
2. Blooming, T. M., & Carnovale, D. J. (2006). Application of IEEE STD 519–1992 harmonic limits. In *Pulp and Paper Industry Technical Conference,* (pp. 1–9). IEEE. https://doi.org/10.1109/papcon.2006.1673767.
3. Machowski, J., Bialek, J. W., & Bumby, J. R. (2020). *Power System Dynamics–Stability and Control*. Wiley.
4. Sookananta, B., Galloway, S., Burt, G. M., & McDonald, J. R. (2006). The placement of FACTs devices in modern electrical network. In *proceedings of the 41st international universities power engineering conference, newcastle-upon-tyne*. UK: IEEE. https://doi.org/10.1109/upec.2006.367585.
5. Ren, H., Watts, D., Mi, Z., & Lu, J. (2009). A review of facts' practical consideration and economic evaluation. In *Asia-Pacific Power and Energy Engineering Conference.* Wuhan, China: IEEE. https://doi.org/10.1109/appeec.2009.4918115.
6. Kusko, A., & Thompson, M. T. (2007). *Power quality in electrical system* (3rd ed.). McGraw-Hill.
7. Martínez, E. B., & Camacho, C. Á. (2017). Technical comparison of FACTs controllers in parallel connection. *Journal of Applied Resources Technology, 15,* 36–44. https://doi.org/10.1016/j.jart.2017.01.001.
8. Kumar, G. V. N., & Chowdary, D. D. (2008). DVR with sliding mode control to alleviate voltage sags on a distribution system for three phase short circuit fault. In *IEEE Region 10 Colloquium and the Third International Conference on Industrial and Information Systems*, (pp. 1–4). Kharagpur, India. https://doi.org/10.1109/iciinfs.2008.4798344.
9. Tümay, M., Teke, A., Bayındır, K. Ç., & Cuma, M. U. (2002) Simulation and modeling of a dynamic voltage restorer. (Vol. 3, pp. 31-35). Adana, Turkey.
10. Dhaked, D. K., & Lalwani, M. (2017). A review paper on a D-FACTs controller: Enhanced power flow controller. *International Journal of Advanced Engineering Technology, 10,* 84–92.
11. Johal, H., & Divan, D. (2007). Design considerations for series-connected distributed FACTs converters. *IEEE Transactions on Industry Applications, 43,* 1609–1618.
12. Kalair, A., Abas, N., Kalair, A. R., Saleem, Z., & Khan, N. (2017). Review of harmonic analysis, modeling and mitigation techniques. *Renewable and Sustainable Energy Reviews, 78,* 1152–1187.
13. Omar, R., & Rahim, N. A. (2009). Mitigation of voltage sags/swells using dynamic voltage restorer (DVR). *Journal of Engineering in Applied Science, 4,* 26–29.

14. Velamuri, S., & Sreejith, S. (2017). Power flow analysis incorporating renewable energy sources and FACTs devices. *International Journal of Renewable Energy Research, 7,* 452–458.
15. Ghosh, A., & Ledwich, G. (2002). *Power quality enhancement using custom power devices.* US: Springer. https://doi.org/10.1007/978-1-4615-1153-3.
16. Hingorani, N. G., & Gyugy, L. I. (1999). *Understanding FACTs: Concepts and Technology of Flexible AC Transmission Systems.* Wiley-IEEE Press.
17. Fardanesh, B., Shperling, B., Uzunovic, E., & Zelingher, S. (2000). Multi-converter FACTs devices: The generalized unified power flow controller (GUPFC). *Power Engineering Society Summer Meeting, IEEE, 2,* 1020–1025. https://doi.org/10.1109/PESS.2000.867513.
18. Wang, F., Duarte, J. L., Hendrix, M. A. M., & Ribeiro, P. F. (2011). Modeling and analysis of grid harmonic distortion impact of aggregated DG inverters. *IEEE Transactions on Power Electronics, 26,* 786–797.
19. Powell, R. (1966). The design of capacitor components of large high voltage AC filter networks. In *IEE High Voltage D.C. Transmission*, (pp. 284–86). Manchester, UK.
20. Noroozian, M., & Taylor, C. W. (2003). Benefits of SVC and STATCOM for electric utility application. In *IEEE PES Transmission and Distribution Conference and Exposition*, (pp. 192–99). USA: Dallas, TX. https://doi.org/10.1109/tdc.2003.1335111.
21. Sharaf, A. M. (1982). Harmonic interference from distribution systems. *IEEE Transactions on Power Apparatus and Systems, 3,* 2975–2981. https://doi.org/10.1109/TPAS.1982.317626.
22. Majstrovic, M. (2003). FACTs-based reactive power compensation of wind energy conversion system. In *IEEE Power Tech Conference Proceedings*, (pp. 41–8). Bologna.
23. Sharaf, A. M., & Gandoman, F. H. (2014). A flexible FACTs based Scheme for smart grids PV-battery storage systems. *International Journal of Distributed Energy Rescue, 10,* 261–271.
24. Sharaf, A. M., & Gandoman, F. H. (2015). A robust FACTs PV-smart grid interface scheme for efficient energy utilization. *International Journal of Power Energy Converters., 6,* 344–358. https://doi.org/10.1504/IJPEC.2015.073614.
25. Sharaf, A. M., & Khaki, B. (2012). FACTs based switched capacitor compensation scheme for smart grid applications. In *International Symposium on Innovations in Intelligent Systems and Applications*, (pp. 1–5). Trabzon, Turkey: IEEE Xplore.
26. Taufik, A. (2013). Search algorithms for engineering optimization. In *Technical Open*, Chapter 8. 13. https://doi.org/10.5772/45841.
27. Geem, Z. W., Kim, J. H., & Loganathan, G. V. (2001). A new heuristic optimization algorithm: Harmony search. *Simulation., 76,* 60–68. https://doi.org/10.1177/003754970107600201.
28. Jayabarathi, T., Yazdani, A., & Ramesh, V. (2012). Application of the invasive weed optimization algorithm to economic dispatch problems. *Frontiers Energy, 6*(2012), 255–259.
29. Castillo, C. A., Conde, A., & Shih, M. Y. (2018). Improvement of non-standardized directional over current relay coordination by invasive weed optimization. *Electric Power Systems Research, 157,* 48–58. https://doi.org/10.1016/j.epsr.2017.11.014.
30. Mehrabian, A. R., & Lucas, C. (2006). A novel numerical optimization algorithm inspired from weed colonization. *Ecological Informatics, 1,* 355–366. https://doi.org/10.1016/j.ecoinf.2006.07.003.

Optimisation of Articulated Vehicular RADAR Antenna Array Parameters Using Machine Learning Algorithms

Asif K. Karim and G. A. Shanmugha Sundaram

Abstract The ever-increasing number of RF-based advanced driver assistance systems (ADAS) applications onboard an autonomous vehicle presents with an opportunity for co-locating and sharing the RF front-ends, which in this context refers to the antenna and its associated circuitry of a RADAR. This paper proposes two sparse planar antenna array topologies synthesised for 77 GHz, by replicating a novel subarray, designed using eight isotropic antenna elements. Elements were placed manually to synthesise the novel subarray geometry, while the superarray topologies were constructed by placing subarrays of a particular orientation in a uniform superarray or by placing subarrays with fixed or optimally selected orientations in a sparse planar array, modified genetic algorithm (MGA) was employed to determine only the optimal subarray positions in the first case while the latter used it to choose both position and orientation, subject to the critical constraints dependent on the selected ADAS applications. Ultimately the comparative study of the topologies showing better performance of proposed topologies are discussed.

Keywords Aperture tuning · Medium range RADAR · Long range RADAR · Subarray · RF front-end co-location · Modified genetic algorithms

1 Introduction

The number and types of sensor onboard an automobile has been increasing steadily over the years, which adds on to the kerb weight of the vehicle and triggers a competition for vantage space among the sensors, which has a potential to affect the form and functionality [1, 2]. As a significant number of current advanced driver assistance systems (ADAS) [3] applications employ RADAR [4] based systems for vehicle

A. K. Karim · G. A. Shanmugha Sundaram (✉)
Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: ga_ssundaram@cb.amrita.edu

SIERS Research Laboratory, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_51

localisation and target tracking as part of the guidance navigation and control (GNC) attributes of ADAS, effective reuse and co-location of the RF front end systems ought to facilitate an efficient SWaP-C solution. This would offer the excess mounting space to other sensors and facilitate aperture sharing between ranging and communication operations or between medium and long-range RADAR operations. This points to a possible solution of implementing a sparse planar superarray [5], which demands a larger mounting space. By definition, the fourth and fifth generation autonomous vehicles don't require human drivers, thus rendering the need for a transparent windshield obsolete and thereby providing ample space to mount the superarray with better beam steering and resolution capabilities [6]. The work discussed here aims to synthesise a novel sparse planar subarray, which could be replicated and positioned with and without orientation change by employing MGA, to device candidate superarrays for RF front-end co-location and sharing.

1.1 Related Work

Several previous researches had explored the idea of aperture co-location and sharing, and these have mainly been undertaken in the synthesis of antennas for jointly operating RADAR and communication [7, 8] systems outside the automotive domain, in contrast to sharing attempted between long and medium-range RADAR applications within the domain. Reconfiguring the antenna operational frequency using RF switches facilitates aperture sharing [9–11]. An RF switch in each subarray unit cell enabled the element to be switched between high and low frequency bands [9]. PIN diodes were used to change the parameters of the resonant structures in [11] for frequency reconfiguration. Suitably combined and converted IF modulated RADAR and communication signals were transmitted to different directions using the same carrier frequency by a Time Modulated Array to achieve dual function RADAR communication [12]. This trend is likely to follow in the automotive domain with the increasing demand for frequency spectrum, due to simultaneous use of (a) low latency communication links between vehicle to everything (V2X) [13] and (b) RADAR sensors employed for situational awareness in autonomous vehicles.

Reuse of RF front-end in automotive domain was attempted in two broad ways, by designing a flat-shoulder shaped beampattern [14, 3] or by intelligently switching between separate antenna arrays synthesised for long range and medium range applications. The drawback of the latter is the lower power efficiency, requirement of larger area for the antenna and complexity involved in the baseband signal processing. Xu et al. in [14] realised a flat-shoulder shaped beam using several series fed patch antenna arrays and a substrate-integrated waveguide (SIW) as the network distributing power and phase.

The selection of a suitable antenna array has considered multiple versions designed for long range and medium range applications [15–18]. A low profile microstrip patch antenna had been proposed with a higher detecting range [15], with an on-chip switch that enabled the reconfiguration between short, medium, and long ranges. Two

miniaturised planar microstrip patch arrays with an inter-elemental distance of half wavelength had been the design element [18] that facilitates switching between long and medium range application by an in-chip switch. Dual-frequency [17] antennas were made using tapered slot antennas, to feed spherical Teflon lenses with frequency selective surface which were reflective at 24 GHz and transparent at 77 GHz in the first design and vice versa in the second. In contrast, [16] designed a polarisation selective planar lens enabling the achievement of a narrow beamwidth and high gain for the horizontally polarised waves, catering to the long-range applications and the vertically polarised waves with wider beamwidth and lower gain catering to the medium range applications.

2 Novel Subarray Synthesis

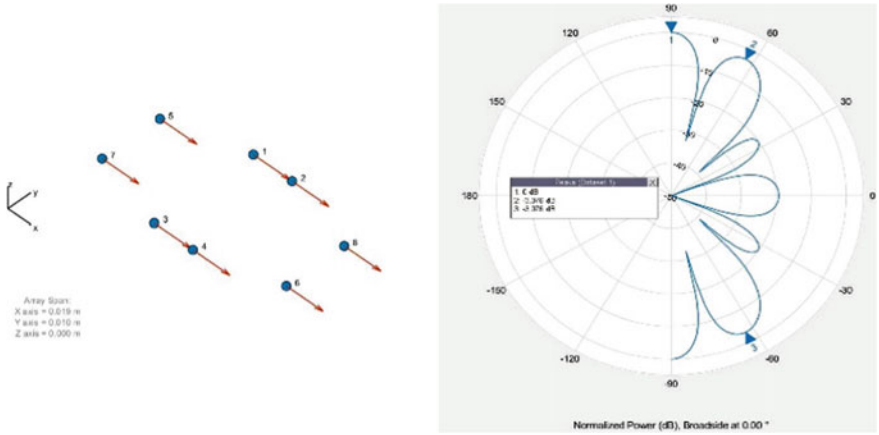
In order to synthesise the novel sparse planar subarray, the antenna aperture was assumed to be on the xy plane with the ordered triplet $(\rho_{xi}, \rho_{yi}, 0)$ representing the position of the ith element of the subarray which could be expressed as $\rho_i = \rho_{xi} + j\rho_{yi}$ in the complex domain such that the $\Re(\rho_i) = \rho_{x_i}$ and $\Im(\rho_i) = \rho_{yi}$. Eight isotropic elements with their elemental normal pointing towards 0° azimuth and elevation were used to devise the reference subarray for 77 GHz.

The subarray elements were fed with uniform power after distributing them on four intertwined sine waves of meticulously chosen frequency, amplitude and phase shifts as in Table 1. The time offset was chosen and the elements were symmetrically located on either side of this point on the sinusoids without violating the minimum separation constraint, $d_{\mathrm{ES}} \geq 0.5 * \lambda$.

The objective of the subarray synthesis was to obtain a symmetric array with a narrow main antenna beam and symmetric side beams, as this could be further tuned to requirement in the superarrays. The synthesized subarray had a dimension of 2×1 cm. ρ is defined as a 3×8 matrix of position vectors of the eight elements in the subarray, post shifting the subarray phase centre to origin. The design of the

Table 1 Depicts details of the four intertwined sinusoids

Sinusoid	Amplitude	Time period (s)	Time offset (s)	Phase shift (rad)
Sinusoid 1	0.0050-	1	$T/4$	0
Sinusoid 2	0▷0500	1	$T/4$	$\pi/2$
Sinusoid 3	0.0050-	1	$T/4$	π
Sinusoid 4	0▷0050	1	$T/4$	$3 \times \pi/2$



(a) (b)

Fig. 1 **a** The subarray geometry. **b** Elevation pattern of the subarray at $\phi = 90°$

subarrays was carried out using the MATLAB Sensor Array Analyzer toolbox, and Fig. 1 shows the geometry of the subarray and the elevation plot.

3 MGA as the Optimisation Algorithm

The optimisation of a sparse planar superarray meant to be vehicle windshield mounted, constituted of N subarrays, allowed to span over a windshield of size $B \times W$ could be achieved by assuming the aperture to be on the xy plane, as illustrated in Fig. 2. Denoting the ordered triplets by (x_i, y_i, z_i) for the coordinates of the subarray phase centres, they could be reduced to an ordered pair (x_i, y_i) as z-coordinate will always be 0, hence represented as a complex value $\omega_i = x_i + jy_i$. Let

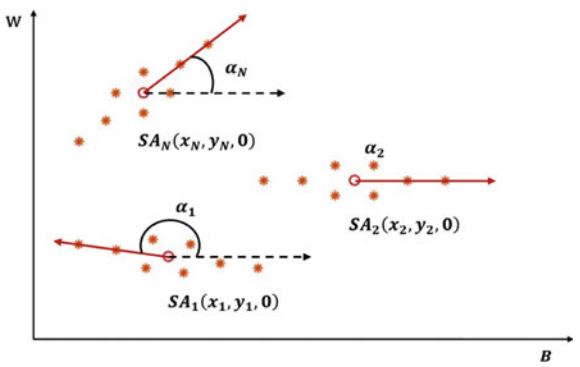


Fig. 2 Geometry of a sparse planar superarray with arbitrarily oriented subarrays

α_i denote the angle in which ith subarray is rotated in the azimuth plane. For optimisation the dimension of the subarray is defined as the radius r of the circle enclosing the farthest element form the subarray phase centre. The aperture considered earlier is assumed to be split into a number of tiles, each tile is expected to accommodate one subarray in the case of a uniform superarray. In that case, the dimension of the tile is chosen in such a way that, when the subarrays are tiled adjacent to each other, an inter-elemental distance not less than d_{ES} is maintained, along with the ability to contain the entire subarray within the tile area for any arbitrary subarray orientation angle.

Square tiles of side $a = 2 \times r + 2 \times 0.5 \times d_{\mathrm{ES}}$ are chosen, $d_{\mathrm{PC}} \geq a$ defines the minimum phase centre separation distance between any two subarrays in the superarray. Now the number of tiles along x and y axes are given by $T_x = \left\lfloor \frac{B}{a} \right\rfloor$ and $T_y = \left\lfloor \frac{W}{a} \right\rfloor$ respectively. Thus, the optimization would be the search for two N dimensional vectors; a complex vector $\Omega = [\omega_1, \omega_2, \omega_3, \ldots \omega_N]^T$ and a real vector $\alpha = [\alpha_1, \alpha_2, \alpha_3, \ldots \alpha_N]^T$, such that the sum of the peak side lobe level (SLL) and the maximum square of error between the expected and achieved half power beamwidth (HPBW) is minimised, subject to the constraints $\frac{a}{2} \leq \Re(\omega_i) \leq B - \frac{a}{2}$, $\frac{a}{2} \leq \Im(\omega_i) \leq W - \frac{a}{2}; i = 1, 2, \ldots N, \alpha_i \in [0, 2\pi]$ and $\max[|x_k - x_l|, |y_k - y_l|] \geq d_{\mathrm{PC}} \geq 0; 1 \leq k, l \leq N; k, l \in Z^+$.

Chebyshev distance which is the sufficient and necessary condition for the fulfilment of Euclidean distance, is chosen as the method of measuring separation between two subarrays in contrast to Euclidean distance to render the non-linear distance constraint linear, thereby reducing the complexity of the GA implementation [6, 19]. Expressions for determining the far-field radiation pattern of the subarray (1) [20] and superarrays (5) [20] were adopted here as the radiation pattern of an array is a function of both its array geometry and constituent elements [21]. The array factor was calculated for an array by replacing $N = N_S$ and $d_n = \rho_i$ in (5) [19]. For calculating the radiation pattern of a superarray having rotated subarrays, the reference subarray is counter clock wise rotated about the z-axis, using the rotation matrix R of corresponding orientation angles, followed by shifting the phase centre of the rotated subarray to the new phase centre location in the superarray. The location of each subarray is obtained as follows $\rho' = R * \rho$; $\rho'' = \Re(\rho') + \Re(\omega_i) + j(\Im(\rho') + \Im(\omega_i))$. The fitness function is defined as in (6b) [19] with an additional term $\left(\mathrm{BW}_{\mathrm{exp}} - \mathrm{BW}_{\mathrm{Ach}}\right)^2$ to optimise the HPBW of the beam pattern.

For GA optimisation, the superarray is assumed to be in positive quadrant of the xy plane as shown in Fig. 2. Hence with an allowance of $0.5 * d_{\mathrm{ES}}$ on either side of the aperture along the x and y axes, the real part of each element in F matrix can take a value in the interval $[0.5 \times a, B - 0.5 \times a]$ and the imaginary part can take a value in the interval $[0.5 \times a, W - 0.5 \times a]$ without violating the aperture constraint. In order to satisfy the minimum element separation distance constraint of $d_{\mathrm{ES}} \geq 0.5 \times \lambda$, the subarrays phase centres should be separated by a distance d_{PC}. Hence, the region left along x-axis after tiling P subarrays is $D_x = B - P \times a$ and along y-axis after tiling of Q subarrays is $D_y = W - Q \times a$. The complex

matrix containing the permissible centroids of the square tiles can be obtained by substituting $d_c = a$ in (11) [19].

The stochastic matrix G and the template matrix F_t are defined as in [19]. In contrast to [19], the elements at the Qth row Pth column is not mandatorily selected when $P\ Q > N$ when populating the F matrix, as it is not necessary to take up the entire available aperture for the case considered in this paper. Now repeat the above process M times to find F matrices for M individuals of the starting population. Let U denote this 3D sparse complex position matrix. A 3D index matrix $S_{Q\times P}$ of M pages, the whole and sparse constraint matrices C_{whole}, C_U were defined, following the definitions in [19]. Post the evaluation of the sparse index matrix, a sparse 2D orientation matrix $\alpha_{Q\times P}$ is initialised corresponding to every page of $S_{Q\times P\times M}$, by populating the fields with valid orientation angles at the corresponding rows and columns wherever there exists non zero values in S. This sparse 3D orientation matrix is named $O_{Q\times P\times M}$. Thus, each of the M individuals represent a feasible solution, which constitutes two 2D matrices, one representing the subarray phase centre positions and the other representing the arbitrary orientation of the subarray in the superarray. After initialisation of the M individuals of the starting population, the fitness values of each of the feasible solutions are evaluated and the population is sorted in the ascending order of fitness.

The crossover and mutation were performed on the chromosome matrix G_F. The two transformations discussed in [19], also have been adopted into this work to ensure that the solution rendered by the optimisation is a feasible one. Genetic pre-processing generates information matrix P [19] and post-processing generates back the offspring population U_C from the information matrix. Throughout the generations, crossover was performed for 25% of the genes of an individual, obtained by genetic pre-processing. This ensures that a better exploring optimisation strategy. On the other hand, mutation was done in a non-uniform fashion with $b = 5$ [22] with a mutation rate linearly diminishing (4) [23] to zero over generations.

The non-uniform mutation ensures that the mutation values lie within the permissible interval. However, in this work, the permissible values of rotation are taken in increments of 0.5° between the lower and upper limits of the interval. Hence, post mutation the angle values were rounded to the nearest midpoint if the values had fractions. However, the subarray phase centres are not rounded to the nearest tile centroid. This will be attempted in a future work. After the completion of crossover and mutation, the real part is sorted along the rows and the imaginary part along the columns, in ascending order without altering the location of the thinned elements. The sorted real and imaginary parts are added back together to get the new information matrix which is genetic post-processed to deliver a feasible solution. The selection of parents for crossover and mutation was done by blending the four equally likely methods (i) Roulette wheel selection (ii) a pair of any two random individuals (iii) best two individuals or (iv) best individual and a random individual pair from a population sorted according to fitness rank, depending on the value of a random integer less than 5 [24].

4 Simulation Results

Superarrays were synthesised for two ADAS applications viz. Adaptive Cruise Control (ACC) and Road Pothole Detection. ACC demands to have a long-range antenna with a narrow beamwidth and pothole detection requires a pencil beam of medium range. Hence, the optimisation targets to achieve an antenna array with beamwidth of 10° for ACC and a beamwidth of 0.87° for pothole detection, while reducing the peak SLL. The superarrays are targeted for a vehicle windshield of dimension 1 × 0.5 m by tiling 20 subarrays, $O_{\text{list}} = \{0.5°, 1.0°, 1.5°, \ldots 179.5°\}$ is the set of orientation angles to choose from, as the novel un-rotated reference subarray is identical to the same rotated by 180°. The minimum elemental separation is chosen as $d_{\text{ES}} = 0.5 \times \lambda$.

Antenna superarrays were synthesised as (i) Uniform superarrays. (ii) sparse planar superarrays with fixed subarray orientation (0°) and (iii) sparse planar superarray with varying subarray orientation followed by a comparison of the beam patterns to determine the best. For achieving this, a program was developed to utilise the Antenna Toolbox and Phased Array System Toolbox in MATLAB.

A 1 × 20 uniform superarray was synthesised by locating the reference subarrays rotated at 35.5° in the counter clock wise (CCW) direction, as 35.5° showed the least value of fitness for the uniform array. The synthesised superarray was 42.6 cm long and 1.6 cm wide, had a HPBW of 11.08° and a normalised peak SLL of −6.71 dB. The sparse planar superarray with fixed orientation was obtained by applying MGA only to achieve subarray phase centre locations, by allowing 100 individuals to reproduce for 1000 generations or until convergence. The optimisation yielded an antenna with a dimension of 96.2 × 46.3 cm, a beamwidth of 1.18°, and a peak SLL of −1.22 dB. Finally, the proposed method acted on 100 individuals for 1000 generations, the optimal solution spanned an area of 90 × 2.5 cm and showed a HPBW of 9.7° with a normalised peak SLL of −14 : 51 dB.

Similar to the above case, a 5 × 4 uniform superarray was synthesised with orientation angle of 83°. The superarray was of size 7.5 × 10.5 cm, with a HPBW of 1.84° and a normalised peak SLL of −5.71 dB. The second method resulted in an antenna with a dimension of 84.7 × 47.3 cm, a HPBW of less than 0.42° and peak SLL of −13.4 dB. The optimal solution derived using the proposed methods synthesised a superarray with HPBW of 0.42° with normalised peak SLL of −9.95 dB. The dimension was 59.5 × 46.4 cm.

By far, the topologies synthesised by the proposed method laid closest to the required HPBW, while keeping the peak SLL at check. The topologies generated by allowing subarray orientation also to change is seen to reduce the size of the array dimension. An extended run of the proposed algorithm was undertaken with 720 individuals for 7200 iterations. The subarray phase centre locations and corresponding orientations in the superarrays have been tabulated in Table 2 for ACC and in Table 3 for pothole detection, the same could be used to obtain various plots of the proposed superarrays.

Table 2 Position and orientation of subarrays in ACC superarray

X	Y	$\alpha°$	X	Y	$\alpha°$	X	Y	$\alpha°$	X	Y	$\alpha°$
0.0109	0.0109	38.5	0.2150	0.0111	32.5	0.3892	0.0107	131.5	0.5672	0.0107	20.0
0.0409	0.0119	160.0	0.2722	0.0107	55.0	0.4154	0.0107	78.5	0.6172	0.0110	146.5
0.0919	0.0115	0.5	0.2996	0.0111	43.5	0.4831	0.0116	112.5	0.6424	0.0109	89.5
0.1137	0.0107	60.5	0.3403	0.0109	86.5	0.5161	0.0107	168.5	0.6944	0.0121	26.0
0.1518	0.0107	150.5	0.3676	0.0107	46.5	0.5377	0.0108	101.0	0.7313	0.0148	6.0

Table 3 Position and orientation of subarrays in superarray for pothole detection

X	Y	$\alpha°$	X	Y	$\alpha°$	X	Y	$\alpha°$	X	Y	$\alpha°$
0.0107	0.0431	24.5	0.1178	0.0139	20.5	0.2249	0.1915	52.0	0.3320	0.0558	23.5
0.0321	0.1560	10.5	0.1392	0.2145	76.5	0.2463	0.2149	11.5	0.3535	0.0496	93.5
0.0536	0.0155	0.0	0.1607	0.0174	48.5	0.2678	0.0141	0.0	0.3749	0.1954	132.5
0.0750	0.2287	62.0	0.1821	0.2418	74.0	0.2892	0.0240	88.5	0.3963	0.2606	134.0
0.0964	0.2112	173.0	0.2035	0.0233	0.0	0.3106	0.0107	111.5	0.4953	0.1860	98.0

5 Conclusion

Out of the three topologies studied in this work reported here, the method of utilising MGA for synthesising sparse planar superarrays, with freedom to change subarray orientations subject to multiple constraints yielded topologies with antenna parameters closest to the design requirements. The main challenge faced by the other two methods is in reducing the peak sidelobe level while tuning the HPBW but the proposal to blend four methods for parent selection ensures that MGA doesn't attain premature convergence. The utilisation of subarrays to populate the superarray, yields a sparse planar array with better beam steering and imaging capabilities and this enables cost saving. Hence, MGA with the changes proposed in this paper may be used to find the optimal locations and orientations for the novel subarray on the sparse planar superarray. The two generated superarray topologies could be arranged on the same aperture and time-shared between targeted ADAS application without jeopardising their critical functioning.

References

1. Abdelhamid, S., Hassanein, H., & Takahara, G., (2014). *Vehicle as a mobile sensor 34*, 286–295, https://doi.org/10.1016/j.procs.2014.07.025.
2. Campbell, S., O'Mahony, N., Krpalcova, L., Riordan, D., Walsh, J., Murphy, A., & Ryan, C. (2018). Sensor technology in autonomous vehicles: A review. In *29th Irish Signals and Systems Conference, ISSC 2018* (pp. 1–4). https://doi.org/10.1109/issc.2018.8585340.

3. Ziebinski, A., Cupek, R., Erdogan, H., & Waechter, S. (2016). A survey of ADAS technologies for the future perspective of sensor fusion. In N. T. Nguyen, L. Iliadis, Y. Manolopoulos, & B. Trawiński (Eds.), *Computational collective intelligence* (pp. 135–146). Cham: Springer International Publishing.
4. Bilik, I., Longman, O., Villeval, S., & Tabrikian, J. (2019). The rise of radar for autonomous vehicles: Signal processing solutions and future research directions. *IEEE Signal Processing Magazine, 36*(5), 20–31.
5. Nie, X. H., Hu, Y., & Hong, W. (2018). Dual-polarized antenna fed by a quasi-planar orthomode transducer with different radiation patterns. *IEEE Transactions on Antennas and Propagation, 66*(11), 5716–5726. https://doi.org/10.1109/tap.2018.2869219.
6. Yu, Y., Hong, W., Zhang, H., Xu, J., & Jiang, Z. H. (2018). Optimization and implementation of siw slot array for both medium- and long-range 77 GHz automotive radar application. *IEEE Transactions on Antennas and Propagation, 66*(7), 3769–3774. https://doi.org/10.1109/TAP.2018.2823911.
7. Feng, Z., Fang, Z., Wei, Z., Chen, X., Quan, Z., & Ji, D. (2020). Joint radar and communication: A survey. *China Communications, 17*(1), 1–27.
8. Masrom, S., Abidin, S. Z. Z., Omar, N., & Nasir, K. (2013). Time-varying mutation in particle swarm optimization. In A. Selamat, N. T. Nguyen, & H. Haron (Eds.), *Intelligent information and database systems* (pp. 31–40). Berlin, Heidelberg: Springer.
9. Haider, N., Tran, D., Roederer, A., & Yarovoy, A. (2011). Reconfigurable L/S band phased array. *Electronics Letters, 47*(23), 1265. https://doi.org/10.1049/el.2011.2899.
10. Sun, F., Zhang, F. S., Zhang, H., Zhang, H., Li, C., & Feng, C. (2018). A frequency diversity printed-yagi antenna element for apertures selectivity wideband array application. *IEEE Transactions on Antennas and Propagation, 66*(10), 5634–5638. https://doi.org/10.1109/TAP.2018.2855722.
11. Xu, J., Hong, W., Zhang, H., Wang, G., Yu, Y., & Hao Jiang, Z. (2017). An array antenna for both long- and medium-range 77 GHz automotive radar applications. *IEEE Transactions on Antennas and Propagation, 65*(12), 7207–7216. https://doi.org/10.1109/tap.2017.2761549, URL https://doi.org/10.1109/TAP.2017.2761549.
12. Shang, X., Zhang, F. S., He, R., & Zhang, H. J. (2019). A 77 GHz miniaturized microstrip antenna array for automotive radar. In *2019 International Conference on Microwave and Millimeter Wave Technology, ICMMT 2019—Proceedings* (pp. 2019–2021). https://doi.org/10.1109/icmmt45702.2019.8992300.
13. Nickel, U. R. (2006). Properties of digital beamforming with subarrays. In *CIE International Conference of Radar Proceedings* (pp. 16–19). https://doi.org/10.1109/icr.2006.343380.
14. Yang, K., Wang, Y., & Tang, H. (2020). A subarray design method for low sidelobe levels. *Progress In Electromagnetics Research Letters, 89,* 45–51. https://doi.org/10.2528/PIERL19110301.
15. Dong, J., & Zhang, L. (2019), Reconfigurable antenna for automotive radar system. In *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting, APSURSI 2019—Proceedings* (pp. 1155–1156). https://doi.org/10.1109/apusncursinrsm.2019.8888847.
16. Schoenlinner, R., Ebling, J. P., Kempel, L. C., & Rebeiz, G. M. (2003). Compact multibeam dual-frequency (24 and 77 GHz) imaging antenna for automotive radars. In *Conference Proceedings—33rd European Microwave Conference, EuMC 2003* (Vol. 2, pp. 785–788). https://doi.org/10.1109/eumc.2003.177593.
17. Syam Sujith, M., & Shanmugha Sundaram, G. A. (2014). SAR image processing using GPU. In *Proceedings of ICCSP 2014* (pp. 866–70). IEEE.
18. Shirazi, M., Huang, J., Li, T., & Gong, X. (2018). A switchable-frequency slot-ring antenna element for designing a reconfigurable array. *IEEE Antennas and Wireless Propagation Letters, 17*(2), 229–233. https://doi.org/10.1109/LAWP.2017.2781463.
19. Chen, K., Yun, X., He, Z., & Han, C. (2007). Synthesis of sparse planar arrays using modified real genetic algorithm. *IEEE Transactions on Antennas and Propagation, 55*(4), 1067–1073. https://doi.org/10.1109/TAP.2007.893375.

20. Liu, F., Masouros, C., Petropulu, A., Griffiths, H., & Hanzo, L. (2020). Joint radar and communication design: applications, state-of-the-art, and the road ahead. *IEEE Transactions on Communications* 1–1. https://doi.org/10.1109/tcomm.2020.2973976, 1906.00789.
21. Cheng, D. K. (1971). Optimization techniques for antenna arrays. *Proceedings of the IEEE, 59*(12), 1664–1674. https://doi.org/10.1109/PROC.1971.8523.
22. Sundaresan, S., & Shanmugha Sundaram, G. A. (2015). Simulation study on modeling the effects of wind turbine on communication signals (C and X bands) using Xgtd. In *Proceedings of ICCSP 2015* (pp. 173–77). IEEE. https://doi.org/10.1109/iccsp.2015.7322867.
23. Michalewicz, Z. (2013). *Genetic Algorithms + Data Structures = Evolution Programs. Springer Berlin Heidelberg.* URL https://books.google.co.in/books?id=jeupCAAAQBAJ.
24. Srivatsa, S., & Sundaram, S. G. A. (2018). PAM4-based RADAR counter-measures in hostile environments. In *Proceedings of 3rd ISTA 2017—Advances in Intelligent Systems and Computing* (Vol. 683(1), pp. 390–400). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-68385-0.

Traffic Congestion Detection from Twitter Using word2vec

Mohammed Ahsan Raza Noori and Ritika Mehra

Abstract Event detection from social media is a highly important field of study. The main goal of the event detection is to detect events (e.g., traffic, disaster) automatically from social media like Twitter. In the area of supervised event detection from Twitter, word2vec model is not yet used by the researchers as a feature extraction technique. Since word2vec introduces a new approach on feature extraction from words and documents, in this work, we propose a near real-time traffic congestion detection system from the Twitter data stream using word2vec model as our feature extraction method. The system fetches the continuous stream of tweets from Twitter, preprocess and extract features from tweets using word2vec, classifies traffic congestion related tweets and notifies the occurrence of traffic congestion in a particular region. For the classification task, we compared the performance of Support Vector Machine (SVM), Logistic Regression (LR) and Naive Bayes (NB) algorithms. Our experimental results have shown that SVM outperforms the other two algorithms achieving an accuracy of 91.73%. Later on, the SVM has been selected to build the actual system.

Keywords Traffic congestion detection · word2vec · Supervised event detection · Twitter · Text mining · Machine learning

1 Introduction

Since the inception of social networks, people use these platforms to communicate real-world events that actually occur around them, and sometimes just share their thoughts on a particular subject by a public tweet. Social media platforms like Twitter have recently been commonly used as a source of information to track events/incidents, like traffic congestion, road accident, building collapse, riots,

M. A. R. Noori (✉) · R. Mehra
DIT University, Dehradun, Uttarakhand, India
e-mail: ahsan.raza@dituniversity.edu.in

R. Mehra
e-mail: hod.mca@dituniversity.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_52

protests, etc. An event can be described as an incident in the real-world which arises within a particular timeframe and place [1]. Specifically, when it comes to traffic-related events like congestion, while commuting people mostly share the current status of traffic around them on Twitter. Traffic congestion is a particular condition when vehicles are moving very slowly, which makes commuters trip times much longer as a result vehicular queuing is increased [2]. Some factors which contribute to traffic congestion are bad roads, narrow roads, increase of personal vehicles, bad weather, traffic violations and road accidents. However, the detection of events from Twitter is more complicated because tweets are unstructured and contain incorrect spelling with grammatical errors. We used machine learning, feature engineering, text mining, and natural language processing (NLP) to overcome these problems and extract meaningful information.

Recent researches have shown that, event detection from social media in terms of feature extraction most of the researchers are using Bag of Words (BOW), Term Frequency–Inverse Document Frequency (TF-IDF), and Parts of Speech tagging for feature extraction. On the other hand, in the text analytics community word2vec is a very popular technique for word embedding as well as for feature extraction [3]. The main aim of this work is to build a near real-time traffic congestion detection system using the Twitter data stream and word2vec as the feature extraction technique. To find a better classification algorithm in conjunction with the word2vec model we have compared the performance of three machine learning algorithms SVM, LR, and NB.

This paper is organized as per the following sections: Sect. 2 deals with work related to different feature representation/extraction techniques used in event detection. Section 3 details the architecture of the proposed traffic congestion detection system. Section 4 discusses the research method used to perform the data collection, preprocessing with feature extraction and classification of tweets. Section 5 lays out the classification results. Section 6 presents the evaluation results of proposed near real-time traffic congestion detection system. Section 7 finally presents conclusion.

2 Related Work

Event detection is a growing area and much noteworthy research has been done to detect events effectively. This literature review is confined to emphasis on different feature extraction techniques used by the researchers in the field of event detection.

Article [4] proposed framework for analysis of Twitter data. For feature extraction Co-occurring terms, TF-IDF and Named Entity Recognition (NER) are used. In [5] the author use the tweet terms to classify the hot topics within a group. In [6] suggests a method for recognizing situational awareness during a catastrophe. The correlation of tweet vectors in word is established using cosine similarity metric. TwitterStand [7] is a news detection system which uses the Term Vector along with some other metadata for the event detection process. Following preprocessing this term vector is produced. [8] identify the event from Twitter dataset and used the combination of features

tweet token, Parts of speech and TF-IDF which gives best performance relative to the individual features and their combinations. Identify the related events in [9], where tweet words are used as features. Stopwords are removed in the preprocessing stage. Tweet terms are used to construct classes of term which is again a collection of specific words. In [10], propose a traffic congestion detection system system on Jakarta, Indonesia road network using TF-IDF as a feature extraction technique. Article [11] focus on detection of traffic events, tokens of tweets are represented in numerical form using IDF score and Information Gain is used for feature selection. [12] develops intelligence reporting framework on arabic social media dataset for disaster management in smart cities. Using text classification, NER, and stemming methods. Features are represented using TF-IDF. On Twitter, a specific research on identifying suicide risk [13] is performed using the TF-IDF value. SVM is used to identify suicidal tweets.

3 Architecture of the Proposed System

Figure 1 shows the proposed system architecture, which mainly composed of 4 modules, namely: (i) Fetching of tweets, (ii) Pre-processing and Feature extraction, (iii) Classification of tweets, (iv) Notification of traffic.

Working of these 4 modules in the proposed traffic congestion detection system are as follows: First, the system fetches a continuous stream of tweets from the Twitter platform using traffic related search query. Second, tweets are preprocessed using text mining techniques like cleaning, tokenization, stopword filtering and feature extraction using word2vec. Third, the system will classify traffic congestion related tweets. Fourth, the notification of traffic congestion as the result of classification is displayed.

Fig. 1 Architecture of the proposed system

4 Research Method

This section describes the various steps carried out to perform data collection, preprocessing with feature extraction and classification.

4.1 Data Collection

We used the Twitter streaming API to capture traffic congestion related tweets. To collect data from Twitter we created a Twitter app to get the API keys and access tokens. These credentials along with a dictionary of traffic congestion related keywords are used to filter the required tweets. Twitter4J,[1] a Java based library has been used to connect and collect the tweets programmatically. The collected data set consists of more than 30,000 tweets from July 2019 to December 2019. Total 30,000 tweets are manually labeled with binary annotation, i.e., as related to traffic congestion (traffic) or not (non-traffic).

4.2 Data Preprocessing and Feature Extraction

Preprocessing is an important step in the classification of texts and has a significant impact on classification performance. Our preprocessing module includes cleaning, tokenization, filtering of stop words and feature extraction.

4.2.1 Cleaning

Even though we have "text" part of the collected tweets, it still contains lots of unnecessary data (@mentions, #hashtags, hyperlinks, punctuations, non-text characters) that we really don't need in our classification process. We have cleaned our dataset to get rid of these unnecessary data by using regular expression filters [14]. Since the #hashtags contain valuable information, so we simply remove the character "#" and keep the rest.

4.2.2 Tokenization

Tokenization is very essential, particularly when it comes to textual data, which is in an unstructured format and actually needs to be transformed into an appropriate form for analysis [15]. Tokenization transforms the documents into a series of words

[1] http://twitter4j.org/.

called tokens and eliminates spaces using the specified separator. In the proposed system, the tokenizer transforms tweets into lower cases and then separates it by blank spaces.

4.2.3 Stop-Word Filtering

Stopwords are by far the most common words in any language which may be of much less importance in the context of analysis. Such words may appear very often in documents but provide absolutely no aid in classification [16]. In the proposed system, the English language stop-word list was freely accessed from the Snowball Tartarus site[2] and supplemented with some other ad hoc stop-words.

4.2.4 Feature Extraction

word2vec represents words by computing its distributed vector. The key benefit of the distributed representations is that similar words are near to each other in the vector space, making generalization easier to novel patterns and more robust model estimation. We used the Skip-gram model for the implementation of word2vec. Skip-gram's training goal is to acquire word vector representations which are excellent at predicting their context in the same document [17]. For a given series of words $w_1, w_2, \ldots, w_T$, the skip-gram model aims to maximize the average log-likelihood

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j=-k}^{j=k}\log p(w_{t+j}|w_t) \tag{1}$$

In the skip-gram model, every word w is associated with two vectors u_w and v_w which are vector representations of w as word and context respectively. The probability of correctly predicting word w_i given word w_j is determined by the softmax model, which is

$$p(w_i|w_j) = \frac{\exp(u_{w_i}^{\top} v_{w_j})}{\sum_{l=1}^{V}\exp(u_l^{\top} v_{w_j})} \tag{2}$$

4.3 Classification

We used three supervised machine learning algorithms, SVM, LR, and NB. These models are evaluated using the 10-folds cross validation process. The metrics used to measure performance of the classification are accuracy, precision, recall, and f-measure.

[2] http://snowball.tartarus.org/algorithms/english/stop.txt.

5 Classification Results and Analysis

The experiments were conducted using Apache Spark 2.4.3 with Scala 2.11 as a programming language. Apache Spark provides different libraries for text mining, feature extraction and machine learning [18]. The experiments were conducted on 64 bit Ubuntu 18.04.4 Operating System with 16 GB memory. Table 1 shows the accuracy, precision, recall, and f-measure for each classifier SVM, LR, and NB. These models are evaluated using the 10-folds cross validation process.

The SVM gives the highest accuracy of 91.73% among the other two algorithms. Figure 2 visualizes the performance comparison chart of SVM, LR, and NB. From Fig. 2, it is noticed that the SVM classifier outperforms LR and NB in all cases. For all metrics values, NB gives the lowest accuracy of 88.68%. The LR gives the second highest accuracy of 90.65%.

6 Detection of Traffic Congestion

To implement the proposed system architecture outlined in Sect. 3, we used the Apache Spark streaming framework [18]. For the detection of road traffic congestion, the developed system was deployed and evaluated for the near real-time tracking of the Indian road networks between January 2020 to March 2020. Table 2 shows some results of traffic congestion detected by our system.

Table 1 Classification results

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
SVM	**91.73**	**91.32**	**91.32**	**91.30**
LR	90.65	90.13	90.12	90.12
NB	88.68	88.42	88.48	88.36

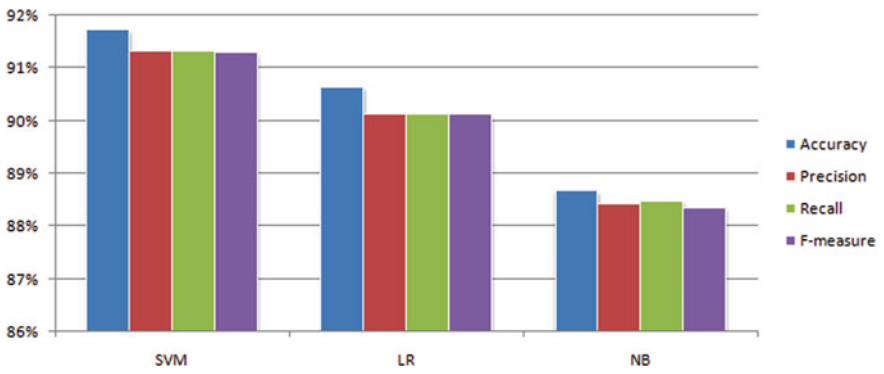


Fig. 2 Comparison chart

Table 2 Results

#	Traffic congestion
1	Heavy traffic at Jaitpur cut, Mathura Road due auto standing
2	There is too much traffic jam, due to pipeline excavation
3	There is a big traffic jam near Juhu signal where express highway joins Juhu
4	Traffic jam in Bavdhan Chandni Chowk near Shinde Petrol Pump
5	Huge traffic build up on Elphinston road in Khadki. No monitoring

7 Conclusion

In this work, we presented a near real-time traffic congestion detection system from the Twitter data stream using word2vec as the feature extraction technique. Three classification algorithms are compared in conjunction with the word2vec model, namely: SVM, LR, and NB. Our experimental results have shown that SVM outperforms the other two algorithms achieving an accuracy of 91.73%. Finally, word2vec model with SVM has been deployed for real-time monitoring of the Indian road networks.

In the future, we aim to enhance and upgrade the system to detect other real-world unusual events in near real-time, like man-made and natural disasters including fire emergencies, building collapse, storm, flood, and earthquake using other popular feature extraction techniques like Doc2Vec, BERT, and GloVe.

References

1. Farzindar, A., & Wael, K. (2015). A survey of techniques for event detection in Twitter. *Computational Intelligence*, *31*, 132–164.
2. Gutierrez, C., Figuerias, P., Oliveira, P., Costa, R., & Jardim-Goncalves, R. (2015). Twitter mining for traffic events detection. In *Proceedings of the 2015 Science and Information Conference. SAI* (pp. 371–378). London, UK.
3. Lilleberg, J., Zhu, Y., & Zhang, Y. (2015). Support vector machines and word2vec for text classification with semantic features. In *Proceedings of 2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing, ICCI*CC 2015* (pp. 136–140). Beijing, China.
4. Gaglio, S., Lo Re, G., & Morana, M. (2016). A framework for real-time Twitter data analysis. *Computer Communications*, *73*, 236–242.
5. Tan, Z., Zhang, P., Tan, J., & Guo, L. (2014). A multi-layer event detection algorithm for detecting global and local hot events in social networks. *Procedia Computer Science*, *29*, 2080–2089.
6. Rogstadius, J., Vukovic, M., Teixeira, C. A., Kostakos, V., Karapanos, E., & Laredo, J. A. (2013). CrisisTracker: Crowdsourced social media curation for disaster awareness. *IBM Journal of Research and Development*, *57*, 1–13.
7. Sankaranarayanan, J., Samet, H., Teitler, B. E., Lieberman, M. D., & Sperling, J. (2009). TwitterStand. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems—GIS 09* (pp. 42–51). Seattle, WA, USA.

8. Ramachandran, D., & Parvathi, R. (2019). Enhanced event detection in Twitter through feature analysis. *International Journal of Information Technology and Web Engineering*, *14*, 1–15.
9. Laylavi, F., Rajabifard, A., & Kalantari, M. (2017). Event relatedness assessment of Twitter messages for emergency response. *Information Processing and Management*, *53*, 266–280.
10. Taufiq, M. (2019). Detection traffic congestion based on Twitter data using machine learning. *Procedia Computer Science, 157*, 118–124.
11. D'Andrea, E., Ducange, P., Lazzerini, B., & Marcelloni, F. (2015). Real-time detection of traffic from Twitter stream analysis. *IEEE Transactions on Intelligent Transportation Systems*, *16*, 2269–2283.
12. Alkhatib, M., El Barachi, M., & Shaalan, K. (2019). An Arabic social media based framework for incidents and events monitoring in smart cities. *Journal of Cleaner Production*, *220*, 771–785.
13. Odea, B., Wan, S., Batterham, P. J., Calear, A. L., Paris, C., & Christensen, H. (2015). Detecting suicidality on Twitter. *Internet Interventions*, *2*, 183–188.
14. Cui, M., Bai, R., Lu, Z., Li, X., Aickelin, U., & Ge, P. (2019). Regular expression based medical text classification using constructive heuristic approach. *IEEE Access*, *7*, 147892–147904.
15. Grover, A., Kumar, P., & Garg, K. (2020). Improvisation of cleaning process on tweets for opinion mining. *International Journal of Big Data and Analytics in Healthcare*, *5*, 49–59.
16. Yao, Z., & Ze-Wen, C. (2011). Research on the construction and filter method of stop-word list in text preprocessing. In *Proceedings—4th International Conference on Intelligent Computation Technology and Automation, ICICTA 2011* (Vol. 1, pp. 217–221). Shenzhen, Guangdong, China.
17. Word2Vec. https://spark.apache.org/docs/latest/mllib-feature-extraction.html.
18. Apache Spark. https://spark.apache.org.

Advanced Signal Processing for Ground-Penetrating Radar

Meghana Nagesha, H. Chetan, and Prasad Borannavar

Abstract Ground-penetrating radar (GPR) is a high-frequency electromagnetic detection device used for land mine detection and other relevant applications such as identification of cement structures, underground metal pieces, and many more fields. This paper addresses the problem of analyzing and improving GPR performance in detecting and characterizing objects in the background of clutter and noise. Significant amount of works in this area have been attempted recently. But, while eliminating direct waves, the traditional time domain method introduces false signals and conflicts and does not provide good random noise elimination. The problem associated with this work is the extraction of the required target signal buried/subdued under noise and clutter whose characteristics keep on changing with the types of soils and external environment. This paper focuses on the enhancement of target information in the background of clutter and noise, by using the simulated data for the comparison and evaluation of different algorithms used for GPR signal processing, against measured data in reduction of clutter and unwanted noise.

Keywords Ground-penetrating radar · Clutter · Noise · Signal processing · Pulse compression · Moving target indicator

M. Nagesha (✉) · H. Chetan
CMR Institute of Technology, Bangalore, KA, India
e-mail: meen16ec@cmrit.ac.in

H. Chetan
e-mail: chetan.h@cmrit.ac.in

P. Borannavar
Electronics & Radar Development Establishment, DRDO, Bangalore, KA, India
e-mail: prasadbgm@lrde.drdo.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_53

1 Introduction

Ground-penetrating radar is an electromagnetic device with high resolution in target identification that primarily aims at buried object identification, earth's shallow surface investigation, construction materials, and landmine detection. GPR provides detailed information on the existence of buried artifacts under favorable conditions. But in case of some specific harsh environments like dusty, stormy, windy, and rainy atmosphere or different types of soil along with thick grass, it affects the performance of GPR particularly in case of land mine detection. GPR transmits an EM wave and awaits a response. A receiving antenna records the variations in the return signal. The resulting measurements recorded during the turn on period of the receiver are called trace. The idealized trace for this simple case consists of a direct pulse and a single reflection from the layer. The receiver output needs to be analyzed whether at a given time the output represents the echo from a reflecting object or represents only noise. The performance of the detection algorithm degrades when the clutters are formed. The clutter occupies a broad spectrum around the target spectrum. The spectral spread either increases the false alarm rate or masks the actual target detection due to a rise in noise depending upon the clutter characteristics. Mathematically, this can be presented as follows. The transmitted EM radar signal is given by

$$E_t(t) = E_0(t_0)\cos[2\pi f_0 t + \phi_0(t)] \tag{1}$$

where:

E_0—amplitude of the signal, f_0—RF frequency of the signal, and φ_0—signal phase.

The received signal is given by

$$E_r(t) \approx \alpha E_0(t - t_d)\cos[2\pi(f_0 + f_D)\cdot(t - t_d) + \phi(t - t_d)] + \text{noise}(t) + \text{clutter}(t) \tag{2}$$

where

T_d—Time delay, α—amplitude reduction, f_D—Doppler frequency, and φ—phase modulation.

The main objectives of this proposed work are the improvement of SNR and enhancement of target information in the background of clutter and noise. This work also focuses on the design and implementing simulated data for the comparison and evaluation against measured data in reducing the amount of clutter and unwanted signals presented in GPR data.

2 Design Philosophy

The methodology proposed in this work falls under piecewise linear analysis, i.e., each algorithm has been developed first and tested with all filters, all phases, and all kinds of Doppler shifts, and the performance checks have been carried out independently. Then, the so-developed model has been inserted along with the required signal processing chain, and the output has been displayed. The results are compared with the system parameters like SNR, range resolution, etc. In this research work, a new technique to solve the problem in-phase clutter and interference is proposed. It consists of the following stages: like data generation, pulse compression, moving target indicator, coherent integration, and power estimation (Fig. 1).

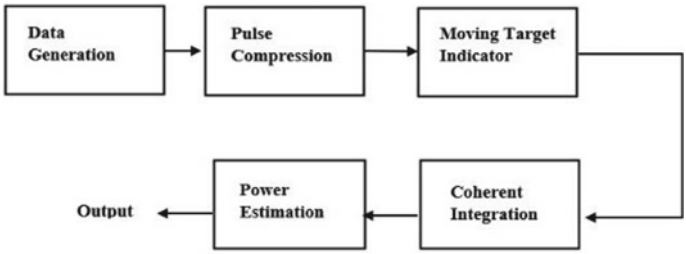


Fig. 1 Proposed block diagram for signal processing of GPR

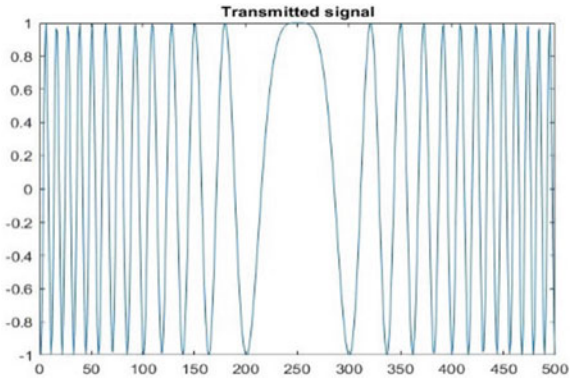


Fig. 2 Simulated LFM signal

3 Methodology

3.1 Data Generation

Linear frequency modulation is known for proving excellent performance when used in pulse compression. Since the LFM spectrum has sharp band edges, it has a better performance in signal processing chains that employ filtering techniques. The LFM waveform data are in complex form and contain in-phase and quadrature phase [1]. The instantaneous phase of the LFM waveform is given by

$$\varphi(t) = 2\pi\left(f_0 t + \frac{\mu}{2}t^2\right) \quad -\frac{\tau}{2} \le t \le \frac{\tau}{2} \tag{3}$$

where f_0 is the radar center frequency and $\mu = (2\pi B)/\tau$ is the LFM coefficient. Thus, the instantaneous frequency is

$$f(t) = \frac{1}{2\pi}\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = f_0 + \mu t \quad -\frac{\tau}{2} \le t \le \frac{\tau}{2} \tag{4}$$

In this work, the bandwidth and radar center frequency have been opted to give the optimal results. The LFM signal is used as the transmitted signal (Fig. 2). For the received signal, a single target along with two in-phase clutters are simulated along with noise as shown in Fig. 3.

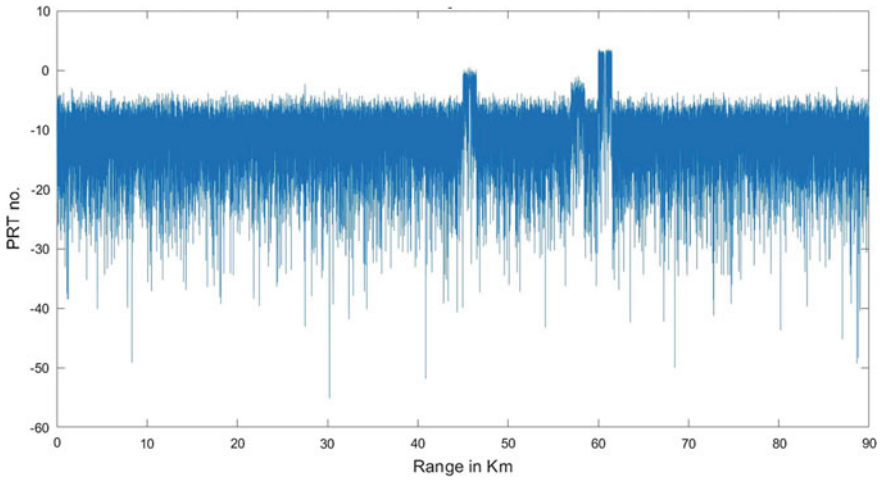


Fig. 3 Received signal

3.2 Pulse Compression

Pulse compression is employed to enhance the target detection capability amid the noise and clutter interference along with increase in SNR. It is a process of modifying the pulse, so that the frequency of the pulse modulation is increased. When the short pulse with wider spectral bandwidth is transmitted, the received radar pulse-echo is passed through a matched filter, which is a low pass filter [2, 3] Thereby, the faster end frequencies are piled up on top of lower frequencies. This results in a higher return pulse output and a narrower pulse width. Thus, pulse compression provides excellent range resolution while packing enough power to provide maximum range. Here, the reference signal is multiplied with a weighted function, and a complex conjugate of this signal is obtained. This is then matched with the received signal to obtain a matched filter output. The frequency domain is chosen for the convenience of performing the correlation. The output of this is then converted back to the time domain by taking the inverse FFT on the output. The output of a matched filter is given by the convolution of the signal $h(\tau)$ with the conjugate impulse response $h * (-\tau)$ of the matched filter [4].

$$Y(t) = \int_{-\infty}^{\infty} h(\tau) h * (t - \tau) \mathrm{d}\tau \tag{5}$$

The impulse response of the matched filter is given by

$$H(\omega) = K S^*(\omega) e^{-j\omega t_0} \tag{6}$$

where K-constant, t_0—delay through filter. $S * (\omega)$—complex conjugate input signal $s(t)$

The matched filter output in the frequency domain is given by (Fig. 4),

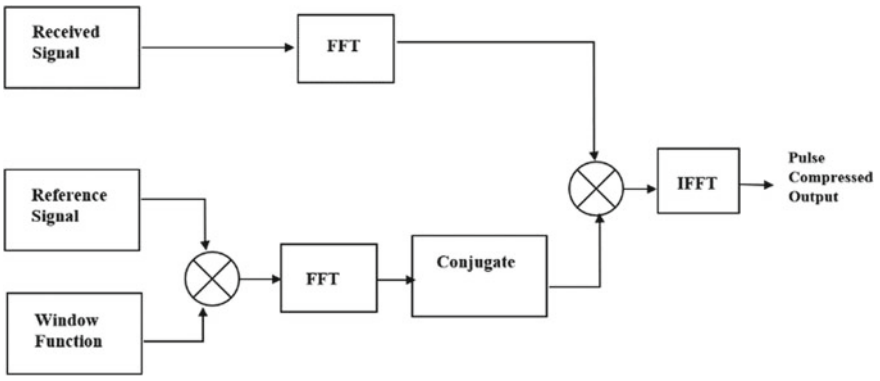


Fig. 4 Block diagram for pulse compression

$$G(\omega) = H(\omega)S(\omega) \tag{7}$$

3.3 Moving Target Indicator

MTI process uses a linear filter to suppress the clutter component in slow-moving time data sequence. Moving target indicator systems are usually operated at relatively low pulse repetition rates to provide an unambiguous range with ambiguous Doppler. MTI usually employs a delay line circuit in its configuration. The working of MTI depends on the fact that the return from the stationary targets having the same phase comparison cancel out, meanwhile the return from moving targets having different phase comparison are displayed. In our proposed work, three configurations of MTI are tested. The three configurations used are single delay line canceller (two-pulse MTI canceller), double delay line canceller (three-pulse MTI canceller), and third-order delay line canceller.

Single Delay Line Canceller
A single delay line canceller has one delay unit, where time delay T is the inverse of the number of pulse-repetitive intervals ($1/f_r$). The received signal $x(n)$, along with the delayed signal, is passed through a summer circuit to get the desired output. The disadvantage in a single delay line canceller is that it produces a very low computational load and has less precision for clutter suppression [5]. The output signal $y(t)$ can be given by

$$y(t) = x(t) - x(T - t) \tag{8}$$

The impulse response is given by

$$h(t) = \delta(t) - \delta(T - t) \tag{9}$$

Double Delay Line Canceller
To overcome the problem of the single delay line canceller, two single delay line cancellers are cascaded to form a double delay line canceller. The clutter rejection notches can be widened through this cascading procedure. The impulse response for this can be given by

$$h(t) = \delta(t) - 2\delta(t - T) + \delta(t - 2T) \tag{10}$$

Third-order Delay Line Canceller
The third-order delay line canceller configuration is formed by cascading a double delay line canceller with a single delay line canceller. This configuration is used to check if it can provide a better clutter rejection ratio than a double delay line canceller.

3.4 Coherent Integration

In pulse compression, the received signal is resolved only in range; therefore, the coherent integration algorithm is needed to resolve the signal. The coherent integration suppresses the noise and interference, thus increasing the signal- to-noise ratio [6]. Coherent integration preserves the phase relationship between received pulses, thus builds up in the signal amplitude . The amplitude of the integrated signal component is increased by a factor of N, with the result that signal energy increased by N^2. The integration is done for all the received pulse-repetitive intervals (PRI). The N-point-FFT to select depends on the number of PRIs. If the number of PRIs are not integral multiples of 2, it is required to do pre-zero and post-zero padding to fill up the input FFT points. In coherent integration, the matrix performs column-wise FFT. Coherent integration gain is equal to the number of pulses coherently integrated.

3.5 Power Estimation

The noise samples, whose phases are varied randomly, are added on a power basis. Power estimation is done to check the final SNR produced. Here, all radar echoes available from the target is summed up in a particular way, and estimated power is used for SNR calculations. It is done by using equation

$$\text{Power Estimation} = 10 * \log_{10}\left(I^2 + Q^2\right) \tag{11}$$

4 Simulated Results

The simulation of the results is carried out in MATLAB. A quantitative analysis has been carried out, and a comparison is given. The LFM signal is generated with a constant pulse duration of 10 μs and center frequency of 250 MHz (Fig. 2).

The received signal serves as an input to the pulse compressor. The pulse compressed output in Fig. 6 shows an enhancement in the amplitude of the targets and clutters while compared to that in Fig. 3. Pulse compressed output shows a greater target visibility with reduced spectral spread. The outputs can be appreciated better in mesh format representation as shown in Fig. 7.

The outputs of single delay line canceller MTI, double delay line canceller, and third-order delay line canceller outputs are shown in Fig. 8. It is quite evident that the stationary clutters have been eliminated at this point. To check the performance efficiency of the delay line cancellers, coherent integration is performed on each of the output.

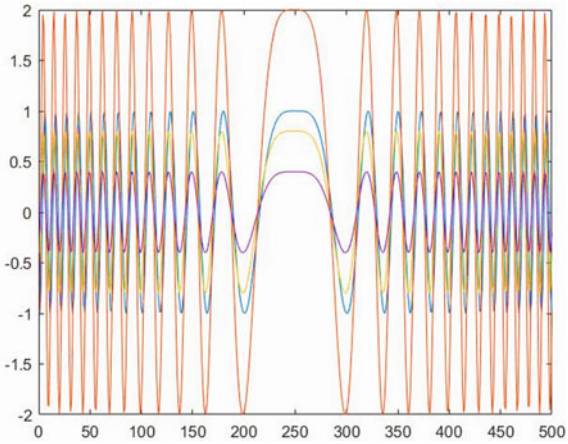


Fig. 5 Phase coherency of target and clutters along with the transmitter signal

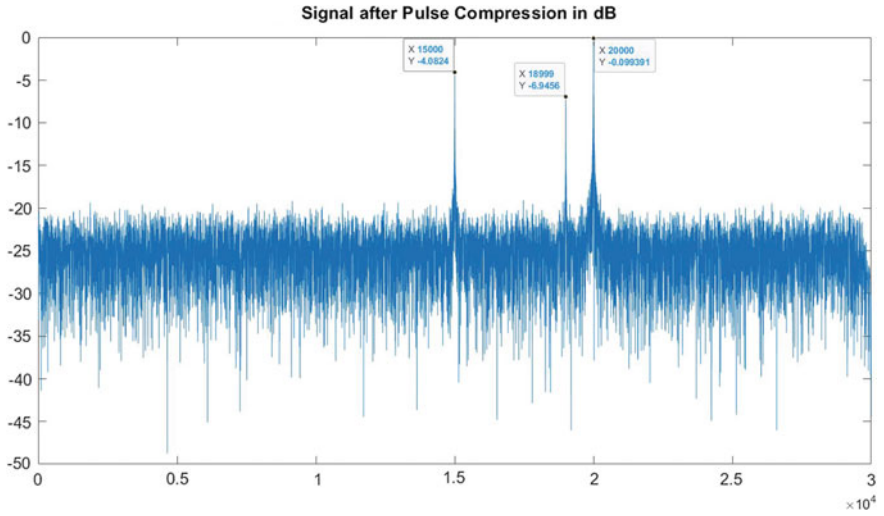


Fig. 6 Output after pulse compression with increased amplitude and SNR

Coherent integration output of a single delay line canceller is shown in Fig. 9. As seen, it has narrow clutter rejection null spectrum. Pulse compression with a single line delay canceller gives a gain of **32.16 dB** in FFT side lobe level.

The coherent integration output with double delay line canceller function is shown in Fig. 10. This configuration indicates a stronger clutter rejection ratio because the highly correlated clutter corresponds to a narrow power spectrum. Thus, high fraction of the clutter energy falls within the filter notch at zero Doppler shift [7].

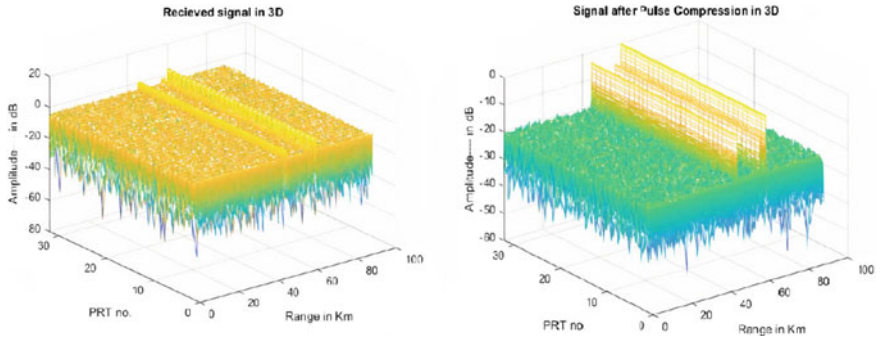


Fig. 7 Mesh outputs before and after pulse compression

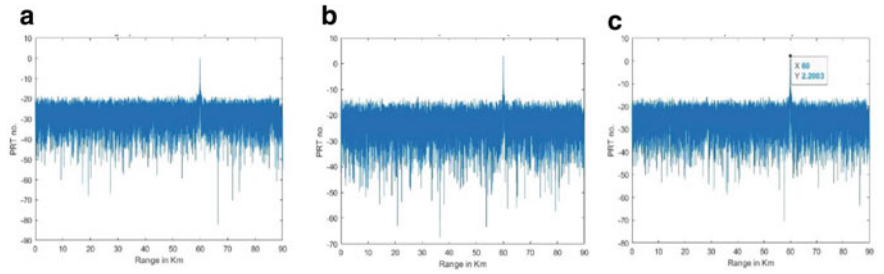
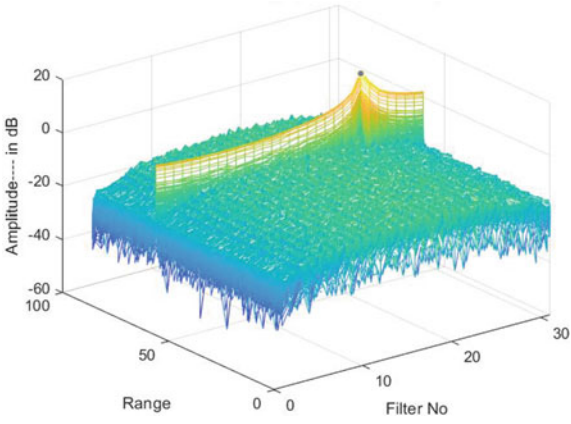


Fig. 8 Delay line cancellers output **a** single, **b** double, **c** third order

Fig. 9 CI output of a single delay line canceller



Pulse compression with double line delay canceller gives a gain of **35.98 dB** in FFT side lobe level.

It can be inferred that the double delay line configuration has a better performance because as the noise level is increasing, the output is seen to take a dip, thereby

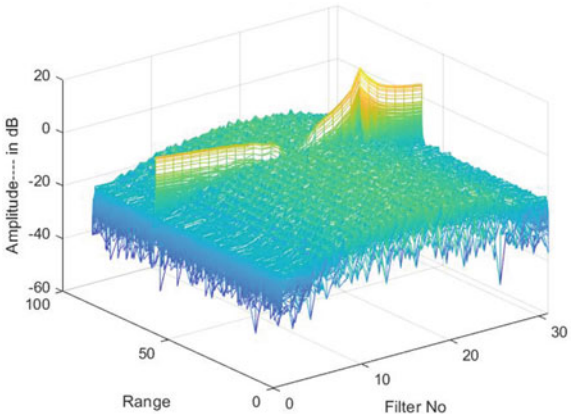


Fig. 10 CI output of double delay line canceller

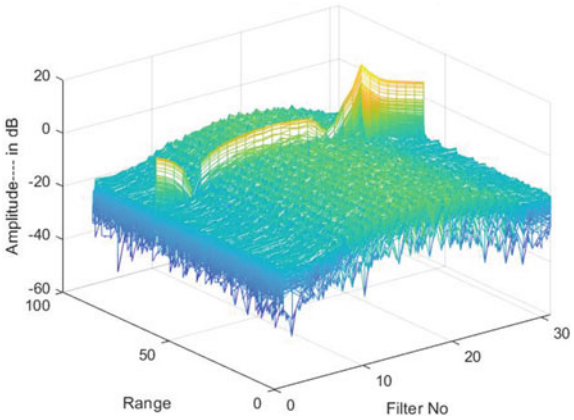


Fig. 11 Third-order delay line canceller CI output

enhancing target visibility. Pulse compression with third-order line delay canceller uses **38.475 dB** gain in FFT side lobe level (Fig. 11). Further enhancement in the order of delay line canceller increases the processing complexity, and also the side lobe ratio is considerably decreased. Hence, a further increase in the order of delay line was not employed.

A comparison of the FFTs with single, double, and triple delay line canceller MTI is shown in Fig. 12. It can be inferred that the third-order delay line canceller provides the best performance when compared to the other configurations.

5 Conclusion

In this paper, an efficient radar detection method is presented which can be used in present day GPRs. The data has been modeled, simulated, and tested using MATLAB,

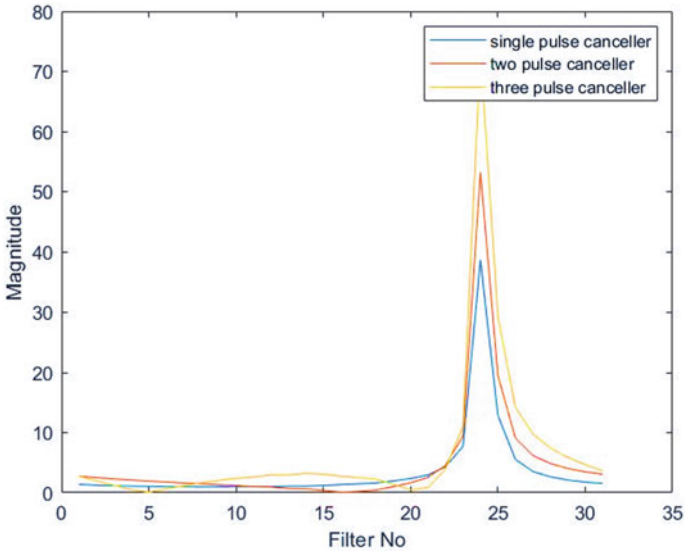


Fig. 12 Comparison of the coherent integration outputs of three configurations of MTI

and the results clearly show the enhancement of SNR. The proposed method increases the dynamic range and range resolution along with reduced computational complexity. The limitations of this scheme are the dynamic usage of pulse cancellers. If one wants effective results, the pulse canceller selection should be done based on the type of soil, where the GPR is employed. Further improvement can be done by employing machine learning techniques for the segregation of the datasets, based on the type of the soil, where the GPR is employed. The decision as to which type of pulse canceller has to be employed can be derived from the machine learning model.

References

1. Ispir, M., & Candan, C. (2016). On the design of staggered moving target indicator filters. *IET Radar Sonar Navig*, *10*(1).
2. Mar, T. T., & Mon, S. S. Y. (2014). Pulse compression method for radar signal processing. *IJSEA*, *3*(2).
3. Archana, M., & Gnana Priya, M. (2014). Low power LFM pulse compression RADAR with side lobe suppression. *IJAREEIE, 3*(7).
4. Skolnik, M. I. *Introduction to radar systems.* New York: McGrawHill.
5. Richards, M. A. (2005). *Fundamentals of radar signal processing* (2nd ed.). New York: McGraw-Hill.
6. Richards, M. A. (2010). *Principles of modern radar volume 1.* SciTech Publishing.
7. Gaballah, M., Sato, M. *A new approach to enhancement of ground penetrating radar target signals by pulse compression.* Clayton: CSIRO Publishing.

8. Parwana, S., Kumar, S. (2015). Analysis of LFM and NLFM radar waveforms and their performance analysis. *IRJET*, *2*(2).
9. Patel, K. (2012). Linear frequency modulation waveform synthesis. *IEEE Students Conference on Electrical Electronics and Computer Science.*
10. Aye1, M. M., ThandarAung, T. (2011). Digital filters for radar signal processing. *IJRET*, *3*(11).

Identification of Intra-abdominal Organs Using Deep Learning Techniques

Sanjeevakumar M. Hatture and Nagaveni Kadakol

Abstract Presently deep learning techniques are playing important role in making healthcare systems more intelligent, efficient, and effective. Proposed methodology is an advisory system in medical imaging which helps in clinical diagnosis. In medical imaging, ultrasound imaging is most frequently used as it is safe, painless, not exposed to ionizing radiation, and it allows real-time imaging. Ultrasound imaging takes more time in diagnosis and well-trained radiologist for interpreting and understanding. Hence, proposed system acts as an advisory system in identifying intra-abdominal organs and abnormalities if any. In this proposed system, the data was collected from intra-abdominal ultrasound images that do not contain any exploring information about the patient. Using filters, noise in ultrasound images is removed. Organ is segmented from ultrasound image and is identified by using deep neural network, and using shape and texture features, abnormalities are identified if any. At the end, various challenges that exist with deep neural network and ultrasound images are discussed.

Keywords Intra-abdominal organs · Deep learning · Ultrasound image · Feature extraction · Segmentation · Region of interest · Disease · Radiologist

1 Introduction

Nowadays, deep learning techniques show tremendous amount of innovations in various fields like Internet applications, IoT, embedded systems, and in healthcare systems. Medical imaging is one of the applications of healthcare system. Medical imaging plays important role in diagnosing the diseases. Ultrasound image is most frequently used in diagnosing the diseases, because ultrasound image technique is

S. M. Hatture (✉) · N. Kadakol
Basaveshwar Engineering College (Autonomous), Bagalkot, Karnataka 587103, India
e-mail: smhatture@gmail.com

N. Kadakol
e-mail: n_kadakol@rediffmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_54

Table 1 Intra-abdominal organs and diseases

Organ	Disease Identified in USA	Organ	Disease identified in USA
Liver	Hepatitis A	Spleen	Cancers
	Hepatitis B		Liver diseases
	Hepatitis C		Inflammations
	Fatty liver		Trauma
	Cirrhosis	Pancreas	Pancreatitis
	Liver cancer		Pseudocysts
Kidneys	Chronic kidney disease		Cysts
	Kidney stones		Neoplasms
	Polycystic kidney disease	Appendix	Appendicitis
	Urinary tract infections		Fecolith
Gallbladder	Gallstones	Small intestine	Infections
	Cholecystitis		Intestinal cancer
	Choledocholithiasis		Intestinal obstruction
	Acalculous gallbladder disease	Large intestine	Colorectal cancer
	Gallbladder cancer		Crohn's disease
	Gallbladder polyps		Intestinal obstruction
	Gangrene of the gallbladder		Ulcerative colitis
	Abscess of the gallbladder	Stomach	Stomach cancer

safe, no exposure to ionizing radiations, painless, and it supports real-time imaging and economically ultrasound imaging is less expensive, hence proposed system uses ultrasound images. Deep learning techniques are used in proposed system to identify intra-abdominal organs from ultrasound image. Presently, most of the people are suffering with problems related to intra-abdominal organs. Some people does not consult clinicians till serious issue, few may have economical issue to consult doctors and some show laziness, and establishing scanning centers at village level, team of radiologist and technicians requires huge financial investments, hence there is scope to develop an advisory system to identify intra-abdominal organ and abnormalities if present. Intra-abdominal organs and abnormalities that can be identified in ultrasound image are listed in Table 1.

2 Literature Review and Discussion

Zhang et al. [1] have proposed a CNN-based model, that classify fatty liver from the ultrasound images as normal, low grade, moderate, and severe fatty liver, texture features, and gray features of the image are used for classification. Normal liver shows the uniform texture, in fatty liver due to the presence of fat particles, and it

shows non-uniform texture. Zheng et al. [2] have used deep transfer learning techniques for classification of kidneys. Anatomic characteristics of kidneys derived from ultrasound images are potential biomarkers of children with congenital abnormalities of the kidney and urinary tract (CAKUT). Classification performance is best by using combination of transfer learning and conventional imaging features. Liu et al. [3] have described applications of deep learning techniques in various tasks such as classification, detection, and segmentation in ultrasound image analysis. Sudheer Kumar and Shoba Bindu [4] have made a review on image analysis using deep learning techniques; it involves identification, classification, and measuring patterns in medical images. Arora and Mittal [5] have used middle, unsharp, and wiener filter for enhancing the ultrasound images.

Angadi and Hatture [6] have employed soft computing tools like symbolic data analysis and multiclass support vector machine for the person recognition. John and Selvan [7] have described the detection of abnormal masses in ultrasound images using texture and shape features. Nalinakshi et al. [8] have used the local binary pattern features to locate the region of interest viz. face for liveness detection of the user. Gupta and Garg [9] have suggested the method for suppressing the speckle and preserve the edges by combining the results of two filters with the help of homogeneity map (HM). One filter is a detail preserving anisotropic diffusion (DPAD) and another is an optimized Bayesian non-local mean filter (OBNLM). Raghesh Krishnan and Radhakrishnan [10] have presented a computer-based approach to classify ten different kinds of focal and diffused liver disorders using ultrasound images. Diseased portion is isolated from the ultrasound image by applying active contour segmentation technique.

3 Proposed Methodology

In this, a new model for identification of intra-abdominal organs such as kidney, liver, gallbladder, pancreas, spleen, and urinary bladder from ultrasound image and abnormalities if exist is proposed. Proposed methodology involves acquisition of ultrasound image, preprocessing by non-local mean filter and bilateral filter, finding the region of interest (RoI), and construction of the deep neural network for intra-abdominal organ recognition, extraction of features for identification and categorization of abnormality in the organ. Proposed methodology is depicted in Fig. 1.

The description of each phase of the proposed methodology is presented in the following subsections.

(a) **Intra-abdominal ultrasound image acquisition**

The system uses the dataset of ultrasound images. The dataset contains ultrasound images of intra-abdominal organs namely kidney, liver, gallbladder, pancreas, spleen, and urinary bladder. Proposed methodology does not disclose any privacy details

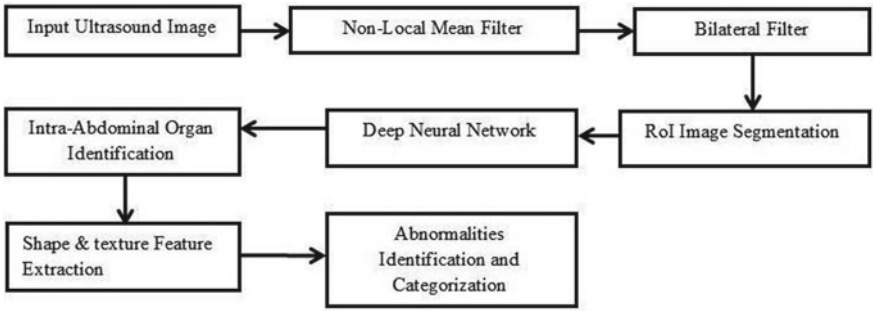


Fig. 1 Proposed methodology

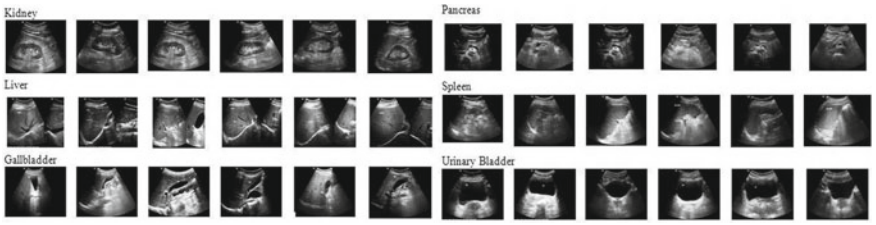


Fig. 2 Ultrasound images of different intra-abdominal organs

pertaining to the enrolled patients as ultrasound images do not contain any information that explores the details like name, age, sex, etc. The sample ultrasound images of dataset are shown in Fig. 2.

(b) **Image enhancement**

Ultrasound images are preprocessed to enhance the intra-abdominal organ information for identification of abnormalities. Non-local mean filter and bilateral filter are used for enhancing the image. The sample of filtered ultrasound images is shown in Fig. 3.

(c) **Segmentation of region of interest image**

Identifying the organ and abnormalities if present, organ is segmented to get region of interest. For segmentation, morphological processing steps are introduced. Figure 4 shows the segmentation of organ. The RoI image of an organ is collected and further used for training the deep neural network. In upcoming section, training the deep neural network is explored.

(d) **Intra-abdominal organ identification using deep neural network**

Pretrained convolutional neural network named GoogLeNet is used for identifying the intra-abdominal organ. GoogLeNet has 22 layers. GoogLeNet classifies images

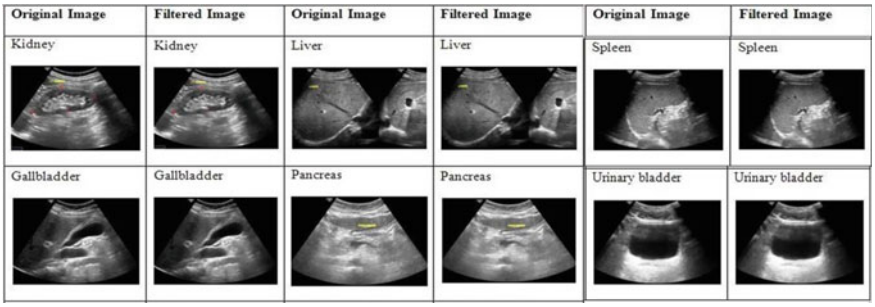


Fig. 3 Original and filtered ultrasound images

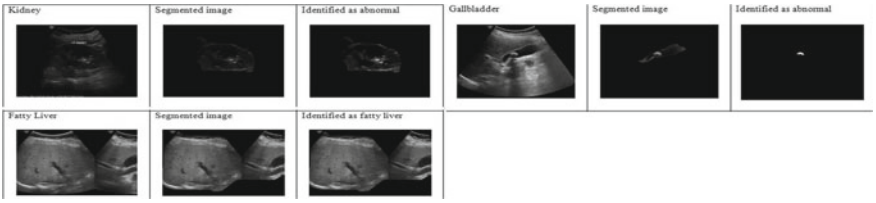


Fig. 4 Shows the original and segmented ultrasound images

into 1000 object categories. Network has an image input size of 224-by-224. In proposed system, GoogLeNet is retrained by using intra-abdominal ultrasound image dataset to classify six categories of organ like kidney, liver, gallbladder, pancreas, spleen, and urinary bladder. After retraining the GoogLeNet network, preprocessed segmented ultrasound image is given as input to the network to classify image as one of the class such as kidney, liver, gallbladder, pancreas, and urinary bladder. Transfer learning approach is used for retraining the network. Retraining the network is illustrated in Fig. 5.

From network, last learnable layer and the final classification layer are used to classify the input image by extracting the image features. 'loss3-Classifier' and 'output' are learnable layer and classification layer in GoogLeNet. In proposed methodology to retrain the network to classify new images, these two layers are replaced with new layers to adapt to the new dataset. Once network is retrained from ultrasound image dataset, 70% of images are used for training the network and 30% of images are used for validating the network.

(e) **Feature extraction**

After the identification of organ, shape, and texture features such as maximally stable extremal regions (MSER), speeded up robust features (SURF), KAZE features, central moment, dispersion, variance, standard deviation and contrast, circularity and solidity are used.

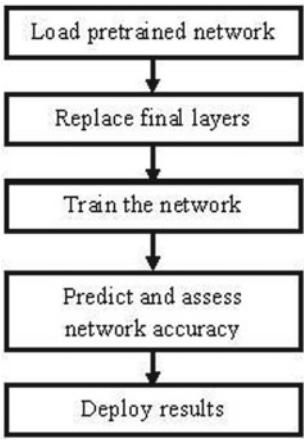


Fig. 5 Retraining the GoogLeNet convolutional neural network

4 Results and Discussion

In proposed method, ultrasound images of kidney, liver, gallbladder, pancreas, spleen, and urinary bladder are used for training the GoogLeNet, and validation accuracy of 98.30% is achieved. Table 2 gives the accuracy of network that correctly identified each organ.

Figure 6 shows the sample of abnormal organs like kidney stones, gallbladder stones, and fatty liver.

Table 2 Performance of the retrained GoogLeNet

Intra-abdominal organ	Accuracy in %	Intra-abdominal organ	Accuracy in %
Kidney	99.54	Pancreas	100
Liver	99.75	Spleen	100
Gallbladder	99.45	Urinary bladder	99.66

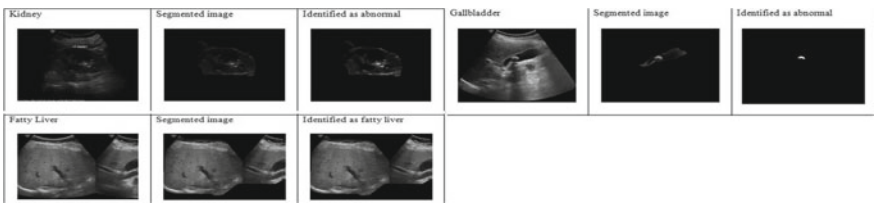


Fig. 6 Kidney stone, gallbladder stone, and fatty liver

5 Future Scope and Perceptive

Based on available set of intra-abdominal ultrasound images, proposed system identifies the organ using deep neural network and abnormalities related to organs like kidney, liver, and gallbladder are identified by using shape and texture features. In kidney and gallbladder, abnormality like stones are identified, and for liver, abnormality like fatty liver is identified. Proposed system of artificial intelligence can be improved and used for various other systems of the body and diseases by collecting extensive data related to more number of patients with different diseases of intra-abdominal ultrasound images and other systems.

6 Conclusion

Proposed system is one of the applications of preventive medicines system which help the clinicians for early diagnosis. Proposed system uses deep learning techniques for identifying organ and by using texture and shape features organ can be classified as normal or abnormalities. Proposed system faces the major challenges like poor signal-to-noise ratio (SNR) and the speckle noise, and soft computing technique such as neural networks requires large training set, partial volume effects, gray-level similarities of adjacent organs, contrast media affect; less availability of data is also a challenging task in automatic identification of intra-abdominal organs and identifying all abnormalities of intra-abdominal organs.

References

1. Zhang, L., Zhu, H., & Yang, T. (2019). Deep neural networks for fatty liver ultrasound images classification. In *Chinese Control and Decision Conference (CCDC)* (pp. 4641–4646). Nanchang, China. https://doi.org/10.1109/ccdc.2019.8833364.
2. Zheng, Q., Furth, S. L., Tasian, G. E., et al. (2019). Computer-aided diagnosis of congenital abnormalities of the kidney and urinary tract in children based on ultrasound imaging data by integrating texture image features and deep transfer learning image features. *Journal of Pediatric Urology, 15*(1), 75.e1–75.e7. https://doi.org/10.1016/j.jpurol.2018.10.020.
3. Liu, S., Wang, Y., & Yang, X., et al. (2019) Deep learning in medical ultrasound analysis: a review. *Engineering*, *5*(2), 261275. https://doi.org/10.1016/j.eng.2018.11.020.
4. Sudheer Kumar, E., Shoba Bindu, C. (2019). Medical image analysis using deep learning: a systematic literature review. In A. Somani, S. Ramakrishna, A. Chaudhary, C. Choudhary, B. Agarwal (Eds.), *Emerging technologies in computer engineering: microservices in big data analytics. ICETCE 2019. Communications in computer and information science* (vol. 985). Singapore: Springer. https://doi.org/10.1007/978-981-13-8300-7_8.
5. Arora, H., Mittal, N. (2019). Image enhancement techniques for gastric diseases detection using ultrasound images. In: *3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 251–256). Coimbatore, India. https://doi.org/10.1109/iceca.2019.8822148.

6. Angadi, S. A., & Hatture, S. M. (2019). Face recognition through symbolic modelling of face graphs and texture. *International Journal of Pattern Recognition and Artificial Intelligence*, *33*(12), 23.
7. John, P., & Selvan, S. (2018). Detection of abnormalities in ultrasound images using texture and shape features. *International Conference on Current Trends towards Converging Technologies (ICCTCT)*, 1–6. https://doi.org/10.1109/icctct.2018.8551174.
8. Nalinakshi, B. G., Hatture S. M., Gabasavalgi M. S. & Karchi R. P. (2013). Liveness detection technique for prevention of spoof attack in face recognition system. *International Journal of Emerging Technology and Advanced Engineering*, 3(12), 627–633.
9. Gupta, M., Garg. A. (2017). An efficient technique for speckle noise reduction in ultrasound images. In: *4th International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 177–180). Noida. https://doi.org/10.1109/spin.2017.8049939.
10. Raghesh Krishnan, K., Radhakrishnan, S. (2017) Hybrid approach to classification of focal and diffused liver disorders using ultrasound images with wavelets and texture features. *IET Image Processing*, *11*(7), 530–538. https://doi.org/10.1049/iet-ipr.2016.1072.

E-Commerce in Indian Retail Industry: Its Proliferation and Performance

Amala Siby and Jossy P. George

Abstract The growth of the e-commerce industry in India has seen a multitude of growth since the growth of netizens in India has reached its peak post the demonetization in Indian economy. Research in e-commerce acts as a catalyst for studies in the field of digital innovation. The developments made by India in the field of e-commerce are notable by the world. India has made extensive use of the advancement in the field of technology. Recent years have seen a transformation in the way Indian shops and exchanges grew from cash mode payments to digital mode of service delivery and payments. This research is focused on studying the parameters that have acted as impetus in the expansion of e-commerce in the Indian retail sector.

Keywords E-commerce · Retailing · Online marketing · Digital advertisements · B2B commerce

1 Introduction

World play projection reports state that India will be the world's second largest e-commerce market by 2034 [1]. The Indian online business market is set to surpass the USA and turn into the second biggest on the planet in less than two decades, dethroning China. The Global Payments Report said that quite a bit of India's Web-based business development will be the consequence of rising Web entrants from 600 million in 2020 to over 1500 million in 2025 [2]. The report also states that with the push for e-governance, the proliferation of smartphones leading to a boom in digital payments due to cheaper data availability all together will act as a catalyst for overall economic growth. As the nation's white-collar class changes over to 3G and 4G systems and Internet being offered at a very cheaper cost, India will see still

A. Siby (✉) · J. P. George
CHRIST (Deemed to be University), Bengaluru, India
e-mail: amala.siby@christuniversity.in

J. P. George
e-mail: frjossy@christuniversity.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_55

more versatile shopping, particularly among millennials. The scope of the Indian retail sector in the e-commerce segment is predicted to have a multifold jump from 15% of total retail sales happening online in the year 2014 to an approximate of 24% in the year 2020.

The Winning Leap, the development of new advances, particularly portable, in India has started a social change that is hard to measure. While versatile, Web and online networking infiltration and development can be measured, portraying the adjustments in social qualities and ways of life that have gone with those patterns is significant all the more difficult. New advancements, for example, virtual dividers, and virtual mirrors will additionally help to enhance the retail client encounter, in this manner empowering more noteworthy utilization [3].

The number of Web clients took off from around 20 million in 2004 to almost above 1500 million in 2020. This fact demonstrates that even uneducated individuals are getting to the Web. While increments in the utilization of customary alternatives for picking up information, seems exponential, the online business industry in India may as of now be behind its partners in various created nations and even some developing markets. Generally, the online business part is developing and various genuine players are entering the market. What separates the Indian Internet business showcase from that of a nation like China is that while advertising focus in China is to a single major player like Alibaba in the market, in India the piece of the pie is partitioned among a few Web-based business organizations, each one thinking of its plan of action.

2 Literature Review

The retail sector in India can be divided into two, structured and unstructured. Traditionally, the Indian retail sector is dominated by several Kirana stores, small-scale departmental stores, etc., which constitutes the unstructured segment of retailers [4]. The structured sector whose size is required to triple by 2022 can be further part up into departmental stores, markets, shopping centers, and so on. Regarding development, the FMCG retail division is the quickest developing unit, and the retail identifying with family unit mind, sweet shop, and so forth have fallen [5]. The remote retail monsters were at first confined from making interests in India [6]. Be that as it may, now FDI of 51% is allowed in India just through single marked retail outlets. Multimark outlets are still past their scope. Again, they can just enter the market through franchisees, this was the way Walmart had entered holding hands with Bharati Enterprises. Online retailing is still to leave a blemish on the clients because of lacunae that we have as of now specified [7]. E-commerce players saw mobile commerce as the most preferred route with a mobile wallet as the preferred way of payment after demonetization. With 4G services launched in 2015, Internet penetration took a significant leap and this trend reached at its peak after Reliance Jio flooded and conquered the telecom industry. This gave another boost to mobile commerce. Changes in lifestyle and shopping choices saw buyers preferring online

Fig. 1 Growth stages of e-commerce in India

and mobile channels over the physical channel to save time and seek wider range and possibly comparative pricing. For mobile wallets, improvements on the payments front with multiple payment instruments and an increase in payment gateways aided by enhanced security with multiple authentication layers helped the consumers with a seamless mobile experience. There has been a clear growth which can be seen when you take the growth of e-commerce in the Indian retail sector [8]. The evolution of growth can be traced from the barter system [9] era to the present scenario of m-commerce as shown in Fig. 1.

More or less we may reason that the retail business in India has a more splendid prospect. It is believed that the revival of this economy can be relied upon to pay and the business era [10].

3 Methodology

To study the growth of e-commerce in the Indian retail industry, seven attributes were found relevant after a detailed review of the literature which has been done. The study revealed the following information related to the contribution of each of these parameters to the growth of e-commerce in the Indian retail industry [11].

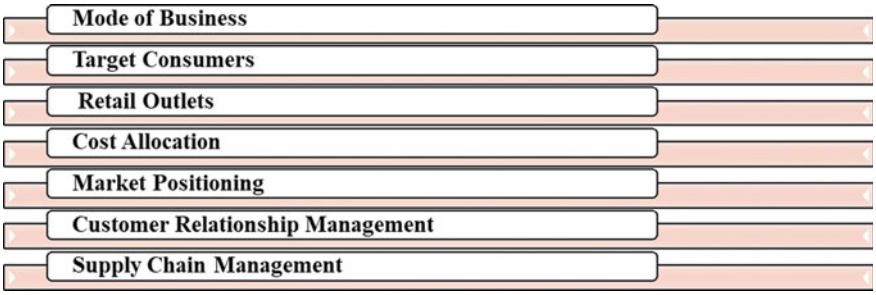


Fig. 2 Attributes of growth of e-commerce in Indian retail industry

The attributes explained in Fig. 2. had a significant role to play when a developing nation like India with an overcrowded population became the potential target group of many multinational companies and conglomerates. These attributes are explained as given below:

3.1 Mode of Business

Since the launch of Make in India campaign by the NDA government on April 25, 2014, there have been constant efforts from the current government to make India into a cashless economy. Demonetization was one among those measures which helped most of the regular customers to shift to digital payments mode [12]. The introduction of unlimited free data by Reliance Jio in 2015 has changed the entire landscape of the situation from the stringent usage of data to unlimited data usage at such low prices [13]. All these factors added up to the shift in business models which earlier focused only on offline business to shift to online modes as well. The traditional business models were replaced by online models to catch hold of the situation.

3.2 Target Consumers

The customer focus was on millennials and the netizen population of the country which helped the business to get more customers. This population was attracted to the convenience that e-commerce has offered concerning product selection, cheaper price, and quality [14]. The growth of such a group gave easy focus for companies to work upon and set the target groups. The growth of digital marketing mainly started in this phase. There are more than 100 million monthly active users for Facebook and in India it is expected that the nation will have the world's largest Facebook population by 2023 [15].

3.3 Retail Outlets

Retail outlets started offering a variety of goods online that enhance the customer base. The convenient shopping concept swept the market [16]. The retail outlets have now modified themselves to be data-driven retail outlets. The introduction of AI-driven tools has been employed by bigger retail players in the country to optimize the performance of retail outlets which has made India as a major influencer of the global retail industry with a faster pace of growth compared to its competitors [10]. Also, the global retail market looks forward to the Indian retail industry as it is a sonorous sector due to its population advantage [16]. To enhance better customer

experience, the retail outlets have tried bringing in changes to the layout of their stores so that visual merchandising using 3D displays gave them a brand-new outlook to attract customers [14].

3.4 Cost Allocation

The cost allocation now started getting distributed among the offline and online businesses equally when the retailers started realizing the acceptability of online shopping among consumers. The attraction of e-commerce to general small-scale retailers was that, though there is huge cost allocation required to establish the Internet presence, it eventually attracted a lot of customers and ensured the online presence of the outlet. There was a shift from e-commerce to m-commerce when companies established themselves in the market and more consumers started getting access to the Internet via mobile phones [10].

3.5 Market Positioning

Every firm started positioning themselves in such a way that they get an impression in the minds of consumers both for their offline store as well as their online presence. This shift of creating an online presence contributed to the growth of many retail stores which earlier were unnoticed.

3.6 Customer Relationship Management

The focus of the business changed to attracting new customers and retaining them. Every business focused on customer retention and considered it as a way for future business growth. A combination of online marketing and mobile marketing has been done to maintain effective customer relationship management [17].

3.7 Supply Chain Management

E-commerce growth has made supply chain management more competitive. Retail business units who have established their presence in e-commerce have now set their focus in delivering quality goods to customers at an affordable price that too by avoiding the time delay. All major e-commerce dealers focus on faster delivery which again boosted the trade happening through the e-commerce industry [1].

4 Results and Discussion

With buyers progressively swinging to online retail, otherwise called electronic retailing, worldwide behemoths are presently competing to catch a bigger share of India's developing Web-based business platforms. Internet business in India represents less than 1 percent of the nations for the most part chaotic retail market and keeps on falling a long way behind that of China and the USA. Following quite a while of moderately moderate development in China, the nation's Web-based business division encountered a blast of action a year ago as Web goliaths Alibaba and Tencent compete for the predominance in the local market. India's populace size and socioeconomics will probably be the main thrust behind the online business' extension soon. While insights demonstrate a considerable lot of India's Web clients do not sign on as often as possible, there is great purpose to trust this will change as the use of cell phones and tablets increases, data becomes cheap, and 4G and growth to 5G broadband scope grow [16].

The pictorial representation of the amount of contribution made by various parameters toward the growth of e-commerce in the Indian retail sector has been represented using percentages in Fig. 3.

The evidence from the compiled pie chart as explained in Fig. 3 has given a clear understanding that the major contributor toward the growth of the e-commerce industry in the Indian retail sector is customer relationship management (CRM). This is major because the introduction of online business has made it appealing to the customers to remain consistent with brands and for brands it is an easy opportunity to retain its customers. The efficiency of supply chain management has also acted as the greatest contributor to the growth of e-commerce in the Indian retail industry by a timely supply of goods at the convenience of consumers. The modes of business have changed from brick and mortar to e-commerce and mobile commerce. A shift in focus on target consumers to a younger population increased cost allocation to

Fig. 3 Percentage contribution of attributes to the growth of e-commerce in Indian retail sector

set up online platforms, and improvement in retail outlet management has together contributed toward the development of e-commerce in the Indian retail industry.

5 Conclusion

The inference from study gives an insight to the brilliant prospect of retail business in India with the existing growth of youth population in the country. If more aggressive promotions strategies are adopted by low costs firms through effective usage of social media platforms, it is found that, the market structure reaction to this strategy is in such a way that it could widen the coverage of minimal cost firms in the retail market turning them out to be more productive. High-cost firms are hit in this process of market expansion by low-cost firms as they employ cost efficient methods to capture the market and customers do not tend to discriminate if the quality is ensured. Social media thus can be said to have brought parity in the promotion mix by economizing the costs involved in making the products and services reach consumers in retail market. Thus, Internet business is assuming an inexorably essential part in Indian retail market. The further scope of research can be extended to the study on the influence of fintech in the Indian retail sector and the implication it has on the development of retail industry.

References

1. Habeeb, S., & Sudhakar, K. (2019). Analyzing Causality Among the Service Quality, Customer Satisfaction and Behavioral Intention Variables with Respect to E-Shopping. *International Journal of Online Marketing, 9*(1), 38–59.
2. Pankaj, M. (2014). *E-commerce Platform in Upcoming Decades.* Bengaluru: Economic Times.
3. R. Boardman, C. E. Henninger and A. Zhu, "Augmented Reality and Virtual Reality: New Drivers for Fashion Retail", Technology-Driven Sustainability, pp. 155–172, 2019.
4. Chattopadhyay, P. (2020). Emergence & Relevance of Organized Retailing in India: A Literature Review. *Tathapi, 19*(06), 281–290.
5. J. Chawla, R. Agrawal and B. Sharma, "A Comparative Study of Unorganized and Organized Retailing in India", Journal of Emerging Technologies and Innovative Research (JETIR), pp. 127–132, 2019.
6. Kamble, A., & Walvekar, S. (2017). Policy regulations in E-commerce sector– Critical analysis of FDI guidelines for market place model. *Journal of Commerce and Management Thought, 8*(3), 409.
7. Prashar, S., Tata, S., Parsad, C., Banerjee, A., Sahakari, N., & Chatterjee, S. (2019). Clustering E-Shoppers on the Basis of Shopping Values and Web Characteristics. *Journal of Global Information Management, 27*(2), 24–38.
8. S. Gupta, "Government Policy for The Development of FDI in Indian Retail Sector: An Overview", Global Journal for Research Analysis, pp. 1–2, 2019.
9. R. Srivastava, G. Santosh Bagale, "Review of Shift from Commerce to E- Commerce in India", International Journal of Advance and Innovative Research, pp. 305–309, 2019.

10. Ghosh, A., Gupta, S., Dua, A., & Kumar, N. (2020). Security of Cryptocurrencies in Blockchain Technology: State-of-art, Challenges and Future Prospects. *Journal of Network and Computer Applications, 163,* 102635.
11. R. Madhukar, "Future of E-Commerce in India Challenges & Opportunities", International Journal of Applied Research, pp. 646–650, 2015.
12. D. Gupta, M. Bindal, N. Agarwal and K. Kumar, "Traditional Commerce v/s E-commerce and Impact of Demonetization on E-commerce", International Journal of Engineering and Management Research, vol. 8, no. 02, 2018.
13. Gawankar, S., Gunasekaran, A., & Kamble, S. (2019). A study on Investments in the Big Data-driven Supply Chain, Performance Measures and Organisational Performance in Indian Retail 4.0 Context. *International Journal of Production Research, 58*(5), 1574–1593.
14. Ali Ibrahim, "The Effect of E-Commerce Application Service Quality of Customer Loyalty Using Customer Relationship Management Approach", 2020.
15. S. Dey and D. Ghose, "E –Retailing & its associated paradigms: A study in India", The International journal of analytical and experimental modal analysis, 2020.
16. S. Tanaji Mantur and H. Borgaon, "Influence of Visual Merchandising on Impulse Buying Behaviour of Customers in Sports Retail Outlet in Hubli City", Studies in Indian Place Names (SIPN), 2020.
17. Kaur, H., Paruthi, M., Islam, J., & Hollebeek, L. (2020). The role of Brand Community Identification and Reward on Consumer Brand Engagement and Brand Loyalty in Virtual Brand Communities. *Telematics and Informatics, 46,* 101321.

Design of a Robot for Healthcare Assistance in the Times of COVID-19

Aayushi Nainan

Abstract With the world reeling under the COVID-19 pandemic, maintaining distance from fellow humans is the new social interaction norm. During these unprecedented times, the medical fraternity and health workers are in dire need of additional help. Robots have been deployed for quite some time for facilitating industrial and healthcare needs. This paper tries to highlight the gap between affordable robots available for health institutions and the demand for them. The aim of this article is to design a prototype robot that can deliver food and medical trays to COVID-19 patients in hospitals to restrict human interaction and hence preventing the further spread of the disease. Autodesk Fusion 360, a computer-aided design tool, has been extensively used to design the robot.

Keywords Robot · Health care · COVID-19 · Fusion 360

1 Introduction

Robots have been making a profound impact in assisting and simplifying medical procedures and healthcare facilities around the globe. It has already established its prominence in niche applications like sensory prosthetics for amputees, exoskeletons in rehabilitation centers for delivering targeted physical and occupational therapy, microbots for delivering medication to a specific part within the body, and humanoids as assistants to doctors [1]. Technological advances in artificial intelligence (AI) sensor technology, 3D printing, and voice activated devices have further opened avenues for scientists and researchers to find applications of robots in enhancing and monitoring a patient's health outcomes.

With the sudden emergence of the corona virus, the world has come to a standstill. Enforcing lockdown in order to force people into maintaining social distancing and to prevent the large-scale local, national, and international human interaction has inspired and motivated the author to contribute to the cause. This article proposes a

A. Nainan (✉)
Student at Dhirubhai Ambani International School, Mumbai, India
e-mail: nainan.aayushi@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_56

design for a low-cost, remotely controlled robot to assist in delivering food trays to patients in hospital wards while maintaining the healthcare guidelines.

The rest of the paper is organized into various sections. Literature review for applications of robots in healthcare institutions has been extensively explored in Sect. 2, while the tools employed for design of the robot along with the description of the proposed product are explained in Sect. 3. Section 4 deliberates the final design of the robot. The article is concluded with suggestions for future scope in Sect. 5.

2 Literature Review

Robotics in the last few decades has emerged as the most challenging field of research especially in industries and healthcare institutions. Extraordinary evolution in use of robots has been witnessed in assisting health patients, administrators, and healthcare facilities to enhance the well-being of the patients by simplifying intricate surgical procedures [2]. With significant research and developments of robots for health care, however when it comes to actual deployment in the field, few have seen the light of the day. Researchers in [3] have proposed a framework wherein the usability of robot from the service scape perspective can be verified. The authors analyzed the equipment, design, space, and ambience parameters of a nursing robot using a GUI design. In applications of wearable robots like myoelectric hand, authors in [4] evaluated the electromagnetic interference of biosensors in the control mechanism. A walking robot partner was well accepted by anxious individuals in a survey conducted in Japan [5]. They also suggested the use of a humanoid robot as an exercise partner in healthcare scenario. With robots finding applications as a human replacement in healthcare sector, it has eased the lives of elderly with mild cognitive impairment. Deploying the robot for assistance is constrained by the lack of intelligence that the robot can have. Research is gravitating toward development of smart robots which can mimic and understand human behavior and react appropriately. The role and extent of embedded AI powered with deep learning in humanoid was explored in [6] with the authors concluding that it immensely optimized the benefits of cognitive training for elderly individuals suffering from age-related cognitive decline. Results of a study on the perception and acceptance of a socially assistive robot platform "Stevie" in improving the working condition of healthcare workers were explored, conducted, and listed in [7]. Its potential to empower the elderly with independence while also providing them with social companionship was asserted by the authors. With amalgamation of advanced sensors, embedded computing and AI technology, authors in [8] have presented an intense and detailed study of robots and human–robot symbiosis in the areas of medication management, assistance to medical professionals, and its role in revamping the healthcare workflow.

Interdisciplinary research has led to collaboration of AI and knowledge of mechanical/electronics engineering to deploy interactive robot agents in healthcare scenarios. Interactive robot was employed during a therapy intervention and the results indicated that participants showed 3.5 times more trust in the therapy as compared to that

delivered by a human agent [9]. Study of intelligent and autonomous robot which can collaborate with humans in social environment has been explored in [10] due to rapid advancements in the areas of AI and Internet of things. To maintain the stability of the robot while walking in different terrains, the authors in [11] proposed an adaptive fuzzy controller method to correct errors on the joint actuators by analyzing foot motion parameters based on ground conditions. Bluetooth and wireless technology have also found its way in optimizing the intelligence of the robots in controlling the motion of arms and can be utilized in healthcare applications [12].

Armed with the information assembled from the various articles researched and with the motivation derived from the current scenario, the next section details the design considerations of the proposed robot.

3 Robot Design Considerations

Various design considerations were deliberated upon before actually getting down to developing the various components required for designing the proposed robot. Since the application of the robot is for a hospital ward environment, the size of the actual robot needed to be worked out. Two-dimensional sketches of the components were then created before proceeding with creating a three-dimensional model. Modifications to the design and deliberating on the appearance of the final product were the last step in the design consideration. Figure 1 enumerates the steps in designing the robot.

The following ideas were adhered to before proceeding further.

- The final design should be like that of a food rack.
- The robot can be operated using a remote controller.
- The patients' respective food tray will automatically slide out based on the linear movement mechanism as given in Fig. 2.
- The patient can collect the food tray without touching anything else.
- The robot will continue to do the same until all food trays have been delivered to the patients.
- Once the trays have been collected in the same way as well, the trays as well as the delivering robot can be sanitized and disinfected before the next round of meals are circulated.

3.1 Fusion 360

The software used to develop the proposed design is Fusion 360. Autodesk Fusion 360, a CAD, CAM, and CAE software, is a cloud-based designing tool that gives one the opportunity to design, model, and simulate your projects digitally, turning their ideas into reality [13]. Replacing the lengthy process of manual designing with a quick, efficient, and automated process, Fusion 360 aids in drastically reducing

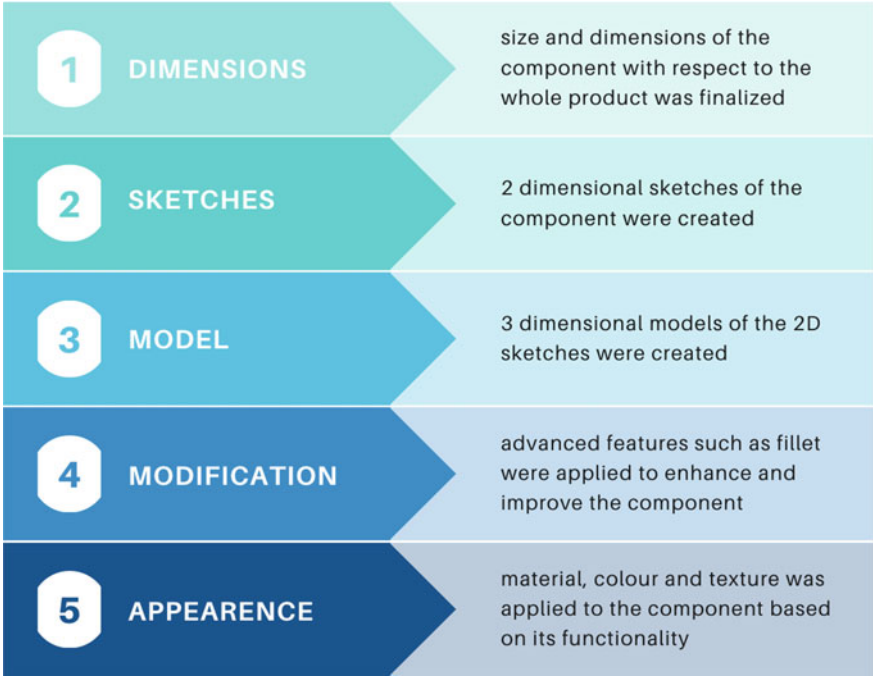


Fig. 1 Steps in implementing design consideration for robot building

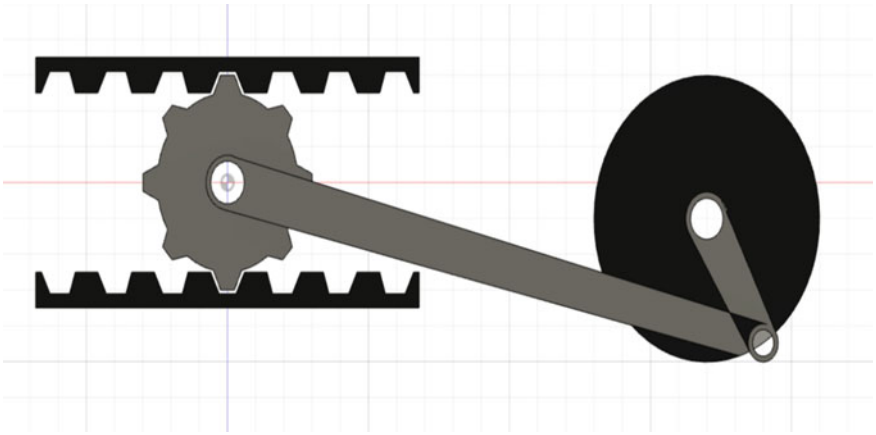


Fig. 2 Linear movement mechanism

manufacturing costs by limiting the need to develop multiple prototypes before finalizing on a design. Additionally, by providing one with three-dimensional views of their designs and enabling them to edit it multiple times until perfection, Fusion 360 produces accurate designs with little room for error. Manually sketching designs and developing different prototypes with each edit results in an overall increase in costs incurred; however, Fusion 360 allows one to make multiple edits to the same design and reduces the need to make prototypes, hence reducing the total expenditure. Additionally, its simulation feature allows one to enhance the product working by indentifying its key structural aspects Additionally, by providing one with three-dimensional views of their designs and enabling them to edit it multiple times until perfection, Fusion 360 produces accurate designs with little room for error. Figure 2 shows the design of a linear motion mechanism that can be implemented in making the physical prototype of the robot.

There are multiple CAD softwares such as Autodesk Inventor, Onshape, TinkerCAD, and Creo. For this article, SolidWorks and Fusion 360 were chosen for comparison. After an extensive comparison between Fusion 360 and SolidWorks, Fusion 360 was the selected software for this article mainly because it provides a large range of geometric tools that can be easily accessed [14, 15]. Furthermore, Fusion 360 offers a variety of robust tools that increase the efficiency of creating 3D models. However, its "multi-component part system" requires one to build all components in the same file in order to assemble it together, which can often be a negative aspect, especially when developing more complex designs [16].

The following Sect. 4 gives a detailed insight into the final design, appearance, and model of each component used to put together the proposed robot. It also contains a brief overview of the features used in Fusion 360 in order to complete each design.

4 Design of the Robot

This section provides a complete visual guide of the different components that were assembled together, creating the final robot. Additionally, the varied features used to create each component has been discussed in Table 1.

The remaining part of this section contains the images of each component separately.

Figure 3 shows the design of the food rack, and Fig. 4 shows the food tray with the food items while Fig. 5 shows the 3D model of the wheel.

After designing each component separately, the final robot was assembled together using various joint features like rigid and revolute as seen above in Fig. 6.

Table 1 Features of Fusion 360 used to design various components of the robot

Component	Features					
	Sketch	Extrude	Revolve	Mirror	Joints	Other
Food rack	Yes	Yes	No	Yes	Yes	–
Food tray + food items	Yes	Yes	Yes	No	Yes	Fillet, appearance
Wheel	Yes	Yes	Yes	Yes	Yes	Line, trim, fillet, appearance
Linear motion mechanism	Yes	Yes	No	Yes	Yes	–

Fig. 3 Design of the food rack

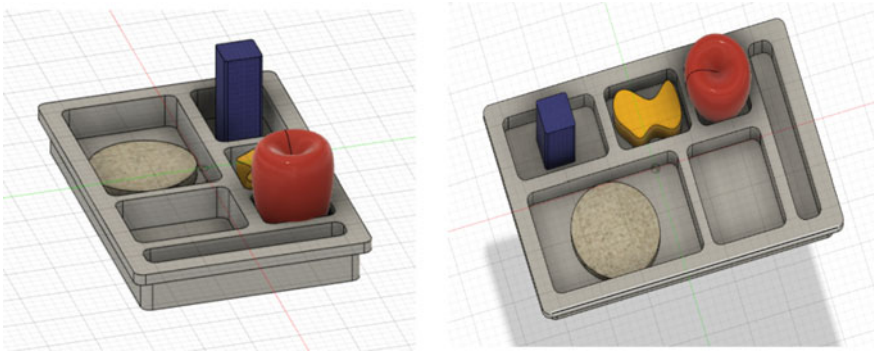


Fig. 4 Design of a sample food tray with sample food items

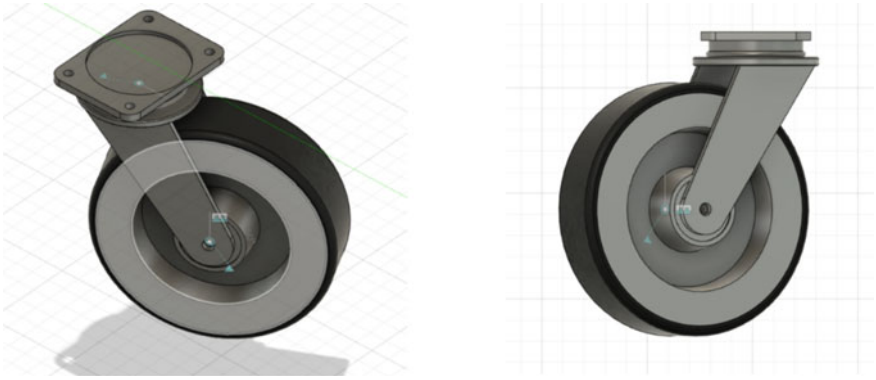
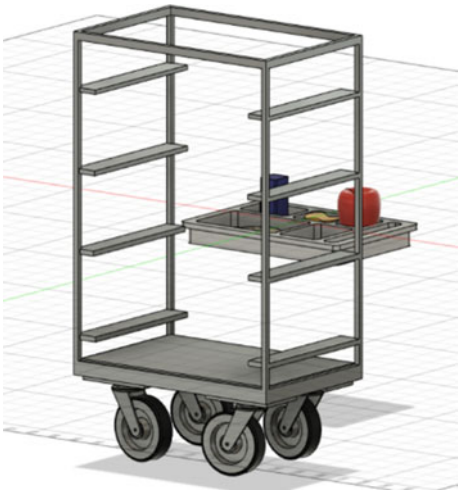


Fig. 5 Design of a three-dimensional wheel using Fusion 360

Fig. 6 Final robot design



5 Conclusion and Future Scope

People from all parts of the world, all walks of life and all age groups are united in their fight against these unprecedented times caused by the corona virus. In an attempt to help, support and assist healthcare workers in hospitals, who are risking their lives to save the infected patients, the author has designed a probable model that can be implemented to protect the frontline workers. According to recent findings, hundreds of healthcare workers have been infected with the contagious coronavirus. With the implementation of the proposed design in this article, the risk of them contracting the disease can be lowered because they will be in less contact with the infected patients.

Although the proposed design has been developed on Fusion 360, the physical prototype is yet to be built. Keeping that in mind, the robot can be implemented on a larger scale, with more food racks to provide more patients, hence increasing in its working efficiency. Currently, the robot can be operated using a remote-controlled device; however, with more advanced research and after its testing, the working of the robot can be automated using the Python programming language.

Additionally, the material used in the design proposed above is aluminum, the same can be implemented using hard acrylic material. While this may make the robot more cost efficient, its ability to remain sturdy and take the weight of multiple food trays is yet to be tested.

References

1. Azeta, J., Bolu, C., Abioye, A., & Oyawale, F. (2018). A review on humanoid robotics in healthcare. *MATEC Web of Conferences*, *153*.
2. Alotaibi, M., & Yamin, M. (2019) Role of robots in healthcare management. In *6th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1311–1314). New Delhi, India.
3. Han, J., Kang, H., Kwon, G. H. (2017). Understanding the service scape of nurse assistive robot: the perspective of healthcare service experience. In *14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)* (pp. 644–649). Jeju.
4. Liao, W., Nagai, K., & Wang, J. (2020). An evaluation method of electromagnetic interference on bio-sensor used for wearable robot control. *IEEE Transactions on Electromagnetic Compatibility*, *62*(1), 36–42.
5. Yamada, S., Nomura, T., & Kanda, T. (2019). Healthcare support by a Humanoid Robot. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 1–2). Daegu, Korea (South).
6. Vogan, A. A., Alnajjar, F., Gochoo, M., & Khalid, S. (2020). Robots, AI, and cognitive training in an era of mass age-related cognitive decline: a systematic review. *IEEE Access*, *8*, 18284–18304.
7. McGinn, C., Bourke, E., Murtagh, A., Donovan, C., & Cullinan, M. F. (2019). Meeting stevie: perceptions of a socially assistive robot by residents and staff in a long-term care facility. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 602–603). Daegu, Korea (South).
8. Pang, V., Yang, G., Khedri, R., & Zhang, Y. (2018). Introduction to the special section: convergence of automation technology, biomedical engineering, and health informatics toward the healthcare 4.0. *IEEE Reviews in Biomedical Engineering*, *11*, 249–259.
9. Xu, J., Bryant, D. G., & Howard, A. (2018). Would you trust a robot therapist? Validating the equivalency of trust in human-robot healthcare scenarios. In *The 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 442–447). Nanjing. https://doi.org/10.1109/roman.2018.8525782.
10. Breazeal, C. (2017). Social robots: from research to commercialization. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction HRI* (pp. 1–1), Vienna.
11. Chen, M., & Chen, K. (2019). Applying a stable dynamic walking mode of humanoid robot based on adaptive fuzzy control. In *IEEE Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS)* (pp. 177–180), Okinawa, Japan. https://doi.org/10.1109/ecbios.2019.8807808.
12. Varshney, V., Kumar, A., Saxena, T., Jha, P., & Tiwari, M. N. (2019). Design and development of intellectual robotic arm. In *6th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 984–987). New Delhi, India.

13. https://academy.autodesk.com/getting-started-fusion-360.
14. https://all3dp.com/2/fusion-360-vs-solidworks-cad-software-compared-side-by-side/.
15. https://total3dprinting.org/solidworks-vs-fusion-360/.
16. https://www.sculpteo.com/blog/2018/10/24/battle-of-software-fusion-360-vs-solidworks/.

A Brief Survey of Challenge–Response Authentication Mechanisms

Prashant Kushwaha, Harshita Sonkar, Fahiem Altaf, and Soumyadev Maity

Abstract Challenge–response authentication mechanism (CRAM) is a procedure to verify the genuineness of a claimant entity and involves the claimant to prove that he/she knows a secret without actually revealing it to the verifier. It involves sending a time-varying challenge by the verifier to the claimant which in turn generates a response by applying a secret-dependent function on the challenge. In this paper, we provide a formal definition of a CRAM and survey research work related to its two types, namely static and dynamic CRAMs. We also analyse the capability of these CRAMs to defend some well-known attacks and also suggest the best practices for the implementation of CRAMs towards the end.

Keywords Challenge–response · Entity authentication · Information security

1 Introduction

Communication is both an integral and vital part of human existence. In today's digital world, because of the tremendous amount of research work over the years, we can communicate with anyone around the globe or in the outer space. However, with invention of powerful tools, the possibility of their misuse comes along. To curb the misuse of communication technologies, several measures must be taken. One of those measures is authentication, which in terms of communication is the way of proving

P. Kushwaha (✉) · H. Sonkar · F. Altaf · S. Maity
Department of Information Technology, Indian Institute of Information Technology Allahabad, Prayagraj, India
e-mail: mit2019004@iiita.ac.in

H. Sonkar
e-mail: mit2019118@iiita.ac.in

F. Altaf
e-mail: pcl2015006@iiita.ac.in

S. Maity
e-mail: soumyadev@iiita.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_57

that someone is genuine or is actually the one who it claims to be. Authentication of an entity is thus a crucial step which if ignored, can lead to information and/or financial loss and its misuse. Although authentication may be achieved in many ways, we will focus on the popular challenge–response authentication mechanism (CRAM) in this paper. Through a better understanding of the CRAM, we will get to know the effectiveness and applications of this mechanism which will help in extending the usage of CRAM in diverse fields related to our daily lives. Some of the important applications of CRAMs include password reset, emergency access to a computer or network, risk-based authentication mechanism in financial and defence services, etc. We begin by explaining the important terminology related to CRAM.

1.1 Terminology

- **Challenge**. It is a question or query that the verifier entity sends to the claimant entity in order to verify the claimant entity's genuineness. The challenge could be static or dynamic. In case of static challenge, the verifier can select the challenge(s) on his own from a set of fixed predefined challenges, e.g. pet name, etc. On the other hand, dynamic challenges are not fixed. They are created based on the belief that there is some information that everyone possesses, e.g. identifying fire hydrants, etc.
- **Response**. It is the answer or solution to the challenge sent by the claimant entity to the verifier entity. It is calculated by the claimant entity with the help of a secret only known to it.
- **Authentication**. It is a process by which the verifier is ensured about the genuineness of claimant.

1.2 Challenge–Response Authentication Mechanism

Challenge–response authentication is a protocol in network security which includes posing a challenge by the verifier from a set of predetermined challenges to the claimant in order to verify its genuineness. If the claimant gives the correct response to the given challenge, then it is authenticated successfully and given access to a computer, network or any other resource. A basic flow diagram of a CRAM is shown in Fig. 1.

In Sect. 2, we present the literature review of different CRAM systems and also discuss their capability to defend against some well-known attacks. In Sect. 3, we provide our suggestions for implementing a secure CRAM. Finally in Sect. 4, we conclude.

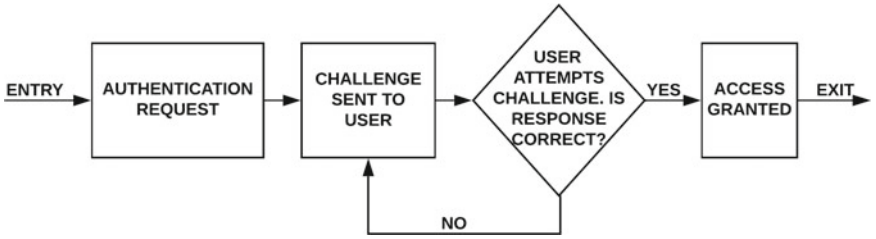


Fig. 1 A basic flow diagram of CRAM

2 Literature Review

The CRAM systems are broadly classified into three types, namely cryptographic, non-cryptographic and hybrid CRAMs. An example of a cryptographic CRAM is the most basic and popular identity/password-based system, while that of a non-cryptographic CRAM is Completely Automated Public Turing Test to Tell Computers and Humans Apart CAPTCHA. Hybrid CRAMs involve both cryptographic and non-cryptographic CRAMs. In this paper, we have focused on pure cryptographic and hybrid CRAMs, both of which involve some form of cryptography. Based upon the type of challenge (as discussed in previous section), CRAMs are either static or dynamic. We will now discuss the related work of these two types of CRAMs below.

2.1 Static CRAMs

- Wang et al. [1] have observed the lightweight stream cipher Trivium of the European eSTREAM project, which is used in Radio Frequency Identification (RFID) tags. They have explored issues like security, privacy, cost and its power consumption. Additionally, different possible attacks on Trivium have also been taken into account. Although it is easy to scan and obtain information from radio frequency identification (RFID) tags, they are not secure, as there is no certification involved. To be economically beneficial, the tags must cost less than \$0.10. At present, the Trivium algorithm consumes 0.68μA@100kHz power with 3488 gates and it has been shown that these figures can be further reduced. Upon further investigation, it was found that the Trivium can be made more secure under the state-recovery attack (breaking complexity under Trivium is $2^{83.5}$). They gave a solution in the form of an improved version of the Trivium algorithm named Enhanced Trivium, which consumes lesser resources and provides the same level of security, by exploiting the internal structure of Trivium. A new set of parameters have been proposed for the algorithm which bring down the power consumption to 0.66μA@100kHz and use only 2876 gates. No attacks have been successful on Trivium till date, and Enhanced Trivium is furthermore secure than Trivium. Spoofing attack has been

countered using a random number, replay attack has been dealt with using an ID, and possibility of desynchronization attack has been curbed by making synchronization mandatory after every five steps. Finally, under state-recovery attack, the breaking complexity of Enhanced Trivium is $2^{86.5}$.

- Jin et al. [2] have studied several reports, e.g. 2018 Internet Security Threat Report [3], threat reports from McAfee [4] which indicate that mobile malware variants, data breaches and fake applications are on the rise. Conventional information and communication security systems for intrusion detection (IDS), intrusion prevention (IPS) and Firewalls have been studied. For obtaining results and testing, a prototype system is created by using a smartphone and a laptop in collaboration with a web server. Already existing Global Positioning System (GPS) information based on a concealed file system (GCFS) is useful, but it is not fruitful in conditions where a GPS is jammed or spoofed. Data backup techniques already exist, but they are not secure and efficient when processing data in external networks. Moreover, information and communication technologies like real-time ransomware monitoring and detection systems, such as IDS, IPS and Firewalls are effective only in internal networks. Virtual private networks (VPN) offer a feasible solution, but they require a good Internet connection. An encrypted QR code-based concealed file mounting authentication method using a challenge–response strategy, involving two mobile devices (e.g. mobile or laptop), has been put forward as a feasible solution. According to the authors, a concealed file system can be built on a mobile system which can be mounted only by another designated mobile device. In this manner, the data will be secure under cyber attacks and also when mobile devices are lost. Moreover, it is suitable in internal as well as external networks, user friendly as well as effective in insecure public networks. Also, it has been successfully demonstrated that a concealed file system can be mounted without network-based communications using a QR-based challenge–response through a smartphone.
- Fujii et al. [5] have highlighted the weaknesses and restrictions of the popular two-factor authentication method based on knowledge factor (ID and password) and possession factor (confirmation on a device). They have studied currently existing authentication technology like bio-metrics in pursuit of a feasible solution. It is well known that an attacker may remotely change the call transfer settings of a genuine user to its own device. Also, it is evident that the caller IDs can be changed in several countries, which makes the solution to this problem more peculiar. Hence, a three-factor method comprising of the knowledge factor, possession factor and additional voice recognition has been proposed by the authors, which can solve the problem with the verification of the possession factor by matching a small audio input sent by the claimant to the verifier. In the method put forward, the user initiates contact with the server for verification of possession factor after the successful verification of the knowledge factor. The server checks the user's caller-ID. To ensure that the caller ID is not forged, there is a provision of an audio challenge, the response to which is a single word (also used for voiceprint recognition). This in turn prevents a replay attack as the challenge is renewed every

time. Despite crossover error rate of voice print authentication on public networks being 6.47%, the authors have successfully defended their proposed method and demonstrated its practicality.

- Rastogi et al. [6] have cited the strengths of the CRAM. A number of password-based cryptographic authentication protocols have been analysed such as the challenge–response mechanism for password exchange, session key exchange methods that offer better security. It has been found that a lot of work has been done on the challenge–response mechanism for password exchange, but a large portion of it is based on assumptions. It has been stated by the authors that most of these solutions were not practically implementable in communication networks. The authors have proposed a dynamic and secure authentication scheme which is based on the unique identification of every client and server pair. Unique list of keys and passwords is generated for every new user for communication. Eavesdropping has been curbed by using a separate challenge for each executable protocol. Encrypted key exchange mechanism confronts dictionary attacks. The private/public key pairs are uniquely and randomly generated to stop masquerading of users.

2.2 Dynamic CRAMs

- Kim et al. [7] conducted a detailed study and analysis of the authentication performance of the conventional physical layer CRAM (PHY-CRAM) specifically with large-scale networks such as Internet of Things (IoT). Several scientific techniques like static hypothesis testing using receiver operating characteristics (ROC) curve, probability density function, rice distributions, cumulative distribution function and Bessel function have been used along with Monte-Carlo Simulations in order to plot and analyse results. Through the simulations and results, authors have shown that the authentication performance of the conventional PHY-CRAM degrades drastically if in case the secret key established between genuine users is 80% compromised, and also the users cannot distinguish between the adversary and legitimate party. Also, false alarm probability (0.8785) for the intrusion signals turns out to be higher in the conventional PHY-CRAM. Hence, they have proposed another PHY-CRAM with the additional feature of dynamic key updation using channel gains between legitimate parties. Through updation of the shared secret key, even if the adversary knows the key, the given mechanism can successfully distinguish the intrusion signals. It has been analysed that the proposed mechanism has better performance for authentication, lower false alarm probability (0.0397), and it can handle impersonation and desynchronization attacks successfully.
- Alharbi et al. [8] discussed several security and privacy concerns related to the rapidly growing IoT. Citing the example of recent reports of Distributed Denial-of-Service (DDoS) attacks on IoT systems, the authors have discussed shortcomings of the proposed cloud-based solutions [9]. Several experiments have been carried out to support the proof-of-concept solutions provided. Cloud duplication is done

using Azure clouds, and DDoS attacks are simulated using software applications. The current solutions for upholding the security and privacy in case of the IoT systems such as cloud computing-based techniques have been shown to be non-scalable due to the latency involved. Also, the present IoT systems are vulnerable to DDoS attacks [10]. The authors have proposed a security system which is based on fog computing and uses VPN for secure access channel and a CRAM to defend the server from DDoS by filtering out suspicious malicious attacks. For this purpose, a decision tree classification technique has been used. The network traffic is classified as trusted, suspicious and untrusted depending on the factors like bursty traffic behaviour, flow count, flow parallelity and flow packet count. In case of suspicious traffic, the source is challenged, thus DDoS attacks can be mitigated. Statistically, it is shown that the proposed solution is 41% better than existing solutions in terms of latency.

- Zhao et al. [11] have discussed about the rapidly growing usage and tremendous capabilities of cloud computing along with its numerous security threats, e.g. authentication, virtual and data security, etc. The critical security and robust privacy requirements have been chalked out in order to put forward an authentication protocol on ad hoc basis. The most common authentication technique involving static passwords has its ups and downs. While simple passwords may be easy to remember, they can be easily broken down and complicated passwords are tough to remember. There are several attacks like shoulder surfing, snooping, sniffing, guessing, etc. The One-Time Password (OTP) has overcome most of these problems, but there is still a counter synchronization problem among client and server. Also, mobile network is a must for OTP authentication, which might not be present at all locations. The authors have proposed a CRAM which is asynchronous in nature, based on smart cards and secret keys. Upon successful application to cloud systems, a multiple element authentication can be achieved and security can be enhanced. Each user is assigned a smart card bearing his identity, USB key and SD key. Then, smart card as well as password authentication is required for login. A third-party authentication centre can be used for the same. Guessing attack is dealt by using one-time token generation. Replay attacks are mitigated by using time factor as challenge whereas hash function ensures integrity. As a result, the overall authentication efficiency and security is enhanced compared to conventional methods.
- Goswami et al. [12] have discussed the weaknesses of the password-based authentication mechanism which is still used in modern-day online applications. Other than that, conventional techniques involving smart cards, tokens, bio-metrics are also considered. The analysis results are derived for the solution presented using the experiments and vulnerability assessment tests, e.g. penetration test, replay attack and use of revoked keys. Concerns like identity theft and e-frauds which are on the rise have been visited. It is found that authentication mechanisms using public key infrastructure (PKI) are not just secure, and they are also accepted by legislation in several nations around the globe, which is the reason they are preferred in administrative and commercial transactions. It has been proved that the

Table 1 Capability of different CRAMs to defend some well-known attacks

CRAMs	IA	DsA	SF	RA	RfA	DdA	GA	Ed	DA	SH
[7]	✓	✓								
[1]		✓	✓	✓						
[8]						✓				
[2]			✓							
[11]				✓			✓			
[5]				✓						
[6]	✓							✓	✓	
[12]	✓			✓				✓		✓

IA impersonation attack; *DsA* desynchronization attack; *SF* spoofing attack; *RA* replay attack; *RfA* reflection attack; *DdA* DDoS attack; *GA* guessing attack; *Ed* eavesdropping; *DA* dictionary attack; *SH* session hijacking

conventional techniques which are based on passwords, smart cards, tokens and bio-metrics do not provide non-repudiation and are subject to loss, theft or duplication. An authentication mechanism using PKI-based authentication technology with integrated cryptographic USB tokens is proposed by the authors. Inclusion of a private key within the token ensures a multiple layer security of the mechanism. A session identifier is used in order to handle replay attacks. The authors have shown that the mechanism can successfully defend against session hijacking, session replay, snooping, eavesdropping, use of stolen keys or impersonating and upholds confidentiality, integrity and non-repudiation.

In Table 1, we have summarized different types of attacks which the above-discussed CRAMs are capable to defend.

3 Suggestions for Secure CRAM Implementation

There are many best practices to be kept in mind while implementing the CRAM systems in institutions or organizations in order to ensure optimum security. They are listed as follows:

- The challenge–response database for all employees or users must be created well in time.
- The challenges whose responses can be guessed or calculated by unauthorized entity must be avoided.
- The challenges must be formulated in a way that the response calculation is near impossible for adversary but piece of cake for genuine user, e.g. easy to remember or calculate, etc.
- Use of personal information with consent of the users can be used in order to create challenges as they are highly unlikely for an adversary to guess.

- At first, the standard questions must be asked followed by personal questions.
- There must be a bound on the number of challenges depending on the security level of the resource to be accessed so that users do not feel fed up.
- The responses should never be displayed, only the successful/unsuccessful verification notification should be generated.
- The accounts of the users should be disabled for a particular amount of time after multiple failed attempts in order to ensure security and conserve network bandwidth and resources.
- Robust encryption techniques must be used in order to protect the challenges.

4 Conclusion

In this paper, we have perused the concept of challenge–response authentication mechanism (CRAM) while discussing its various fundamental concepts and terminology necessary to understand it. We have classified the CRAMs into two types, namely static and dynamic CRAMs. The CRAMs discussed in this paper are either purely cryptographic or involve some form of cryptography along with non-cryptographic procedures to carry out authentication. Besides, we have discussed the attacks, vulnerabilities and their solutions related to various CRAMs we have mentioned in the literature, which will open new dimensions of possibilities for further research and improvements in this field. We have also listed some suggestions which can act as a checklist for organizations to implement secure CRAM systems.

References

1. Wang, S., Zhang, S., Liu, L., & Chen, G. (2016). A challenge-response authentication for RFID using Enhanced-Trivium. In *2016 International Conference on Security of Smart Cities, Industrial Control System and Communications (SSIC)*, Paris, 2016 (pp. 1–5).
2. Jin, Y., & Tomoishi, M. (2019). Encrypted QR code based optical challenge-response authentication by mobile devices for mounting concealed file system. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, Milwaukee, WI, USA, 2019 (pp. 676–681).
3. Security Response Publications. *2018 Internet Security Threat Report* (Vol. 23). https://www.symantec.com/security-center/threat-report. Accessed May 14, 2019.
4. McAfee Labs Threats Report December 2018. https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-dec-2018.pdf. Accessed May 14, 2019.
5. Fujii, H., & Tsuruoka, Y. (2013). Three-factor user authentication method using biometrics challenge response. In A. R. Sadeghi (Ed.), *Financial Cryptography and Data Security*. Lecture Notes in Computer Science (Vol. 7859). Berlin, Heidelberg: Springer.
6. Rastogi, N., Pathak, A., & Rastogi, S. (2013). Enhanced authentication scheme using password integrated challenge response protocol. *International Journal of Computer Applications*.
7. Kim, S., Lee, Y., Choi, J., Jeon, M., & Hwang, E. (2019). Dynamic key update strategy in physical-layer challenge-response authentication. In *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, Australia, 2019 (pp. 1–6).

8. Alharbi, S., Rodriguez, P., Maharaja, R., Iyer, P., Subaschandrabose, N., Ye, Z., et al. (2017). In *IEEE 36th International Performance Computing and Communications Conference (IPCCC)*, San Diego, CA, 2017, (pp. 1–2).
9. Kakanakov, N., & Shopov, M. (2017). Adaptive models for security and data protection in IoT with Cloud technologies. In *Proceedings of IEEE MIPRO'17* (pp. 1001–1004).
10. Yang, Y., Wu, L., Yin, G., Li, L., & Zhao, H. (2017). A survey on security and privacy issues in Internet-of-Things. *IEEE Internet of Things Journal, 99*, 1–10.
11. Zhao, G., Li, Y., Du, L., & Zhao, X. (2015). Asynchronous challenge-response authentication solution based on smart card in cloud environment. In *2015 2nd International Conference on Information Science and Control Engineering*, Shanghai, 2015 (pp. 156–159).
12. Goswami, S., Misra, S., & Mukesh, M. (2014). A replay attack resilient system for PKI based authentication in challenge-response mode for online application. In *2014 3rd International Conference on Eco-friendly Computing and Communication Systems*, Mangalore, 2014 (pp. 144–148).

A Model for Automated Food Logging Through Food Recognition and Attribute Estimation Using Deep Learning

Aditi Ambadkar, Chetana Chaudhari, Megha Ghadage, and Madhuri Bhalekar

Abstract The past few decades have witnessed an increase in dietary ailments, majorly caused due to unhealthy food habits. According to experts, mobile-based diet monitoring and assessment systems can capture real-time images of various food items as well as analyze the nutritional content present in it and can be very convenient to use and assist in improving food habits. This can help people lead a healthier lifestyle. This proposed model provides an innovative system that can automatically estimate various food attributes like the nutrients and ingredients by classifying the food image that is given as input. The approach involves the use of different types of deep learning models for accurate food item identification. Apart from image recognition and analysis, the food ingredients, nutrients and attributes are obtained and estimated by extracting words which are semantically related from a large collection of text, accumulated over the Internet. Experimentation has been performed with the Food-101 dataset. The proposed system assists the user to obtain the nutritional value of the food item in real-time which is effective and simple to use. The proposed system also provides supporting features such as food logging, calorie tracking and healthy recipe recommendations for self-monitoring of the user.

Keywords Food image recognition · Attribute estimation · Vector embeddings · Convolutional neural networks (CNN) · Web scraping · Diet monitoring · Food logging · Calorie counter

A. Ambadkar (✉) · C. Chaudhari · M. Ghadage · M. Bhalekar
School of Computer Engineering and Technology, MIT World Peace University, Pune, Maharashtra, India
e-mail: aditiambadkar@gmail.com

C. Chaudhari
e-mail: chetana.chaudhari24@gmail.com

M. Ghadage
e-mail: meghaghadage123@gmail.com

M. Bhalekar
e-mail: madhuri.bhalekar@mitwpu.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_58

1 Introduction

Food items that are very high in calories and fat content can prove to be fatal to our health and as a consequence of their consumption can lead to obesity. Such food may cause several ailments like heart attack, abnormal blood pressure, diabetes and high cholesterol and other diet-related problems. A study conducted by Lancet [1] conveys that an unhealthy diet is a cause for a greater number of premature deaths that resulted from high blood pressure and tobacco. The second biggest factor that led to deaths in 2017 in India was dietary risks [2]. Diet management is very essential, and the main challenge is to keep a balance between what type of food one eats and how one keeps a track of his/her food habits. The existing applications such as SHealth [3] and MyFitnessPal [4] focus on monitoring the food consumption of users. The drawback of these applications is that the user has to provide the details of the food manually, including its serving size. This process can become very repetitive, tedious and time consuming. There are a few methods that use smartphone cameras to automatically identify the food item [5–9]. Nevertheless, these techniques are unable to measure the food attributes.

The goal of the proposed system is to create awareness and improve health, diet and fitness among the people so that they can take steps toward leading a healthier life. The goal is to automatically identify the food item in the image and then obtain the nutrients such as the amount of carbohydrates and proteins as well as the ingredients and calories contained in the food. The system also consists of calorie tracking and food logging for self-monitoring and suggests healthier recipes for the food item taken into consideration.

Convolutional neural networks (CNNs) [10, 11] are extensively used in object recognition and have proven to be successful; thus, they are used for the identifying food item present in an image [6–9]. Deep neural networks are applied for obtaining ingredients and food attributes such as fry, spicy and grill. The output of the implemented system is a Web application and mobile application named "Eatiza," wherein the user uploads an image of the food item through camera or gallery, identifies it and provides the food attributes, nutrients and ingredients present in it.

1.1 Contributions

The main contribution of this project is the two modules—the food recognition module implemented using CNN and the attribute estimation module developed using vector space representation of words. The latter module is trained on data obtained using Web scrapping from food nutrition and recipe Web sites. Additional functionalities such as food logging are provided for diet management and calorie counting, for keeping a track of one's progress. Web links to healthier recipes of food items are suggested to promote better food habits. Both a Web and a mobile application of the proposed model are implemented.

The following portion of the paper is structured as follows: In Sect. 2, an overview of the related work is provided. Section 3 focuses on methodology details. Section 4 illustrates the results of experimentation and implementation. Section 5 provides the conclusion of the paper.

2 Related Work

In order to understand the effect of mobile applications in healthcare processes [12, 13], several studies have been conducted. Research has also been conducted on how social media can be used for health-related purposes.

In [14], food items are classified into six groups based on their similar nutritional attributes such as grains, vegetables and oils using CNN architecture. However, this system does not identify the food item name and only provides its group. Jiang et al. [15] use deep learning techniques for food recognition and diet assessment. In this system, the feature map for food identification is extracted using deep neural networks.

Yunus et al. [16] explains a system that uses a CNN to train the food recognition engine. The food attributes like nutritional values of food are obtained using vector embeddings. Im2calorie [17] extracts the food classes, ingredients and volume of each food item and calories. However, the dataset which is calorie annotated being employed is insufficient [18]. The main technique for the above-explained methods is to begin with food category identification, estimate the food portion size and lastly estimate the calories using standard nutritional fact tables.

Mikolov et al. [19] introduces methodologies that can be employed for learning word vectors from large vocabulary. It describes various model architectures such as Recurrent Neural Net Language Model (RNNLM), and log-linear models like skip-gram models and continuous bag of words (CBOW). Zhang et al. [20] proposed a food image recognition system employing CNN, which have been used for image recognition successfully. This model used UEC–100 dataset, and it can achieve all the recognition steps including distortion and shift invariance, feature extraction, and classification and then provides the output label.

In [21], system contains neural network that can extract features from food images and obtain calorie quantities from features and NutritionX API in real time. In [22], Thai food images are recognized based on transfer learning of CNN. In [23], food images in Food-11 dataset are classified using pre-trained Inception-v3 CNN model.

3 Methodology

This section focuses on the modules that are integrated to build the proposed system. The two major modules are:

- Food Recognition Module—Given a food image as input, this module recognizes the food item and assigns the corresponding food class label.
- Attribute Estimation Module—This module is responsible for approximating the ingredients, attributes and nutrients of the recognized food item using vector space embeddings.

The proposed system also incorporates support modules to assist in diet management, to track user progress and provide insightful analysis to the user. Food logging facilitates the user to maintain a food diary, which contains daily records of food intake. Maintaining a food diary has proven to improve diet and achieve dietary goals for effectively [24]. The calorie counter determines daily calorie intake goals based on user details. It monitors calorie consumption by considering factors like daily calorie intake, calories burnt during exercise and food intake.

3.1 System Architecture

The system architecture of the application is shown in Fig. 1. The food recognition module is responsible for recognizing the food item. The attribute estimation module provides the nutrients, attributes and ingredients of the food item. The application also provides calorie tracking and food logging facilities.

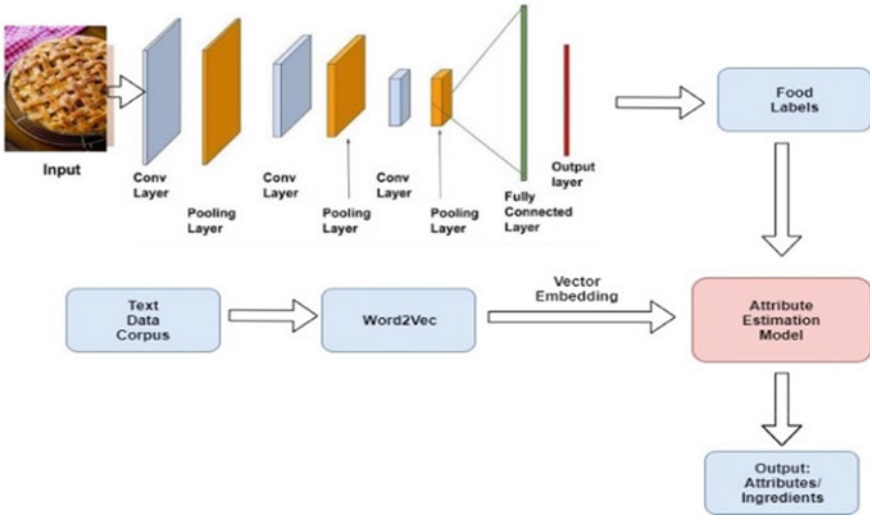


Fig. 1 System architecture of the proposed model

3.2 Dataset

The Food-101 is employed for this purpose. It comprises 100 food categories having 1000 images for each category. Each food image consists of a single food item only. The food item is assumed to be of standard serving size; hence, the nutritional contents are estimated in accordance with the same. The dataset is divided into training and testing datasets. Each food category comprises 800 training images and 200 testing images.

3.3 Food Recognition

This module takes a food image as input, recognizes the food item and outputs the corresponding class label. Experimentation was conducted on the top-performing pre-trained models such as Inception-v3, Inception-v4 and ResNet using transfer learning. It is a CNN that is trained on the ImageNet dataset. The model is trained on the Food-101 dataset by employing transfer learning. The fully connected layer present in the end is taken off and appended with the dropout, ReLU activations and softmax layers.

Fine Tuning. In order to improve the validation accuracy of the model, several techniques are applied. Data augmentation—horizontal and vertical shift and flip operation, random rotation, zoom and standard normalization—is performed. This is done so that the model can adapt to affine variations and the images are trained efficiently. We have performed rescaling by dividing each pixel value by 255, so that they are normalized between the range of zero and one. Batch normalization and regularization are also performed to improve accuracy—avoid overfitting.

3.4 Attribute Estimation

Attribute estimation is the second step—obtaining nutrients, attributes and ingredients of the food item. Figure 2 represents the attribute estimation module. In the proposed system, the vector space model of words from a large data corpus is used. A huge amount of textual data is collected from the Internet using Web scrapping to compute accurate, relevant and precise results. Word2Vec is a tool for creating word embeddings which are used for training the collected data. This module estimates attributes of the food by calculating the distance between the food label and the attributes in the learned vector space. If the distance between the attribute and food item is small, it implies that the probability of the attribute related to that food item is high.

Retrieval of Text Data and its Preprocessing. A Web crawler called scrappy is used to gather data from Google search results. The search is performed with respect

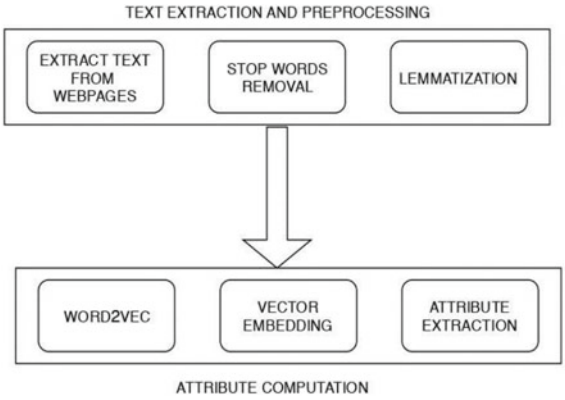


Fig. 2 Attribute estimation module

to each food class, ingredient and attribute on Google and from the resulting pages. Data retrieved from the first 250 pages for every search query is appropriate as it retrieves pages corresponding to precise labels.

The raw text data collected from Google is in HTML format. This has to be preprocessed to remove various HTML tags, JavaScript, CSS code, comments and to obtain individual words by tokenization. Irrelevant and auxiliary words (stop words) such as "the", "is" and "that" are removed for semantic-based text processing. Lemmatization is performed to avoid distribution of probabilities due to various forms of the same word. Thus, the resulting corpus of words is used for training purpose.

Training and Vector Space Representation. The Word2Vec tool employs shallow neural networks such as continuous bag of words (CBOW) and skip-gram model to effectively compute vector embeddings in a multidimensional space. Each word contained in the text data corpus is taken as an input to the log-linear classifier that learns the words which are present within a certain range of input. It establishes semantic relationships by using cosine distance measure between the word vectors. Word2Vec produces a word vocabulary from words used in the text from input data, learns the distances between words and outputs a binary file comprising learned vector space representations.

Attribute Extraction. After a food item is identified, the predicted food class label and static list of ingredients, nutrients and food attributes are provided to the Word2Vec, which finds the cosine distance of the vector embeddings of the food label and the attributes. Similarity function of Word2Vec module returns probability of single words within window of another word using cosine distances. Finally, its probability of existence in the food item is obtained. Since ingredients have the highest frequency of occurrence, they are the most accurate because their sample size is large. Nutritional values are comparatively sparse in the text, whereas food attributes occur more infrequently, hence less accurate. Normalization is performed by dividing class probabilities with the largest probability of that class to get more accurate results.

Table 1 Accuracies of different models for food recognition

Models used	Training accuracy (%)	Testing accuracy (%)
Inception-v3	85.22	72.75
Inception-v4	89.24	88.50
Xception	92.20	86.08
InceptionResNetV2	95.18	92.40

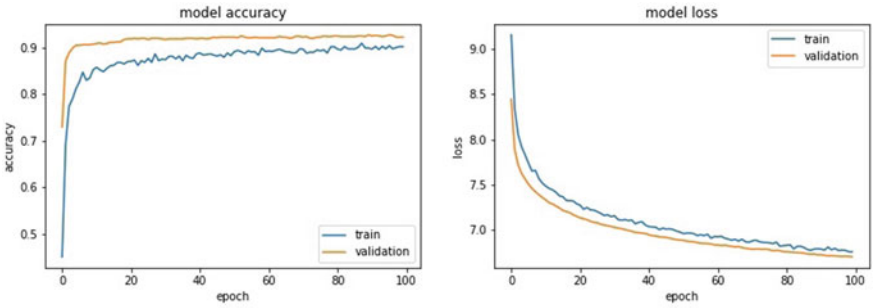


Fig. 3 (i) Training and validation accuracy curves over 100 epochs (ii) training and validation loss curves over 100 epochs

4 Results

4.1 Food Recognition

All images from the dataset are fed to a data augmentation module that generates batches of augmented images which can be later fed to the model. Different pre-trained models are used for experimentation. Top four performance results of the trained models are shown in the Table 1. It can be observed that InceptionResnetV2 shows the highest performance with an accuracy of 92.40%.

Performance of the model over epochs can be analyzed from the following graphs (Fig. 3). The learning curves show improvement in the accuracy and the loss of model over training and validation sets. This helped to identify which model parameters have to be optimized, and the model with the best accuracy and loss is saved.

4.2 Attribute Estimation

Gensim provides pre-built word embedding models like Word2Vec or fastText trained on large corpora. To get accurate results for specialized data, i.e., information related to only food items and nutrition, a new Word2Vec model is trained on generated

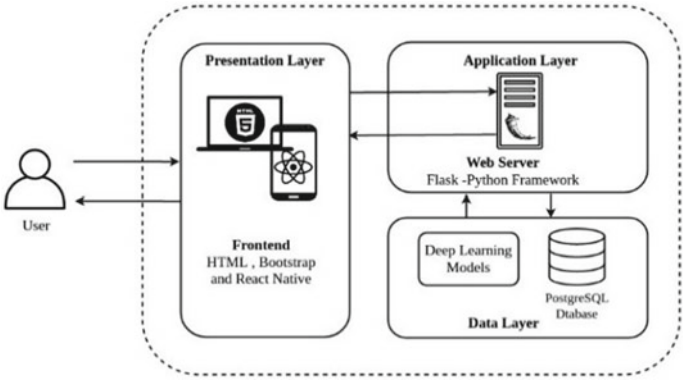


Fig. 4 System architecture of Web and mobile application

Table 2 Parameters for Word2Vec model implementation

Parameter name	Value
Window size	10
Training algorithm (CBOW/Skip-gram)	CBOW
Sampling method	Negative sampling
Subsampling	1e−5
Vector dimensionality	150

word corpus. After testing with different parameters, a model is finalized with the following parameters (Table 2)—CBOW is selected for training as it trains faster than skip-gram and gives better accuracy for frequent words.

4.3 Combined Results

Both the modules, i.e., the food recognition and attribute estimation modules, are combined to produce the final results. Figure 5 shows results obtained for pizza and chicken wings. Our application "Eatiza" uses a three-tier architecture and is implemented on a local machine. System architecture of the application is shown in Fig. 4. Web application is developed using HTML and bootstrap. Mobile app is implemented using React Native: a JavaScript framework. The deep learning models are deployed as an API on a server. The Web server is built by using Flask: Python Web framework, and it uses a PostgreSQL database to store the results and user details.

Ingredient	Value	Nutrient	Value	Food Item	Ingredient	Value	Nutrient	Value	Food Item
Pepperoni	1.0000	Fat	1.0000		Chicken	1.0000	Protein	1.0000	
Cheese	0.9556	Protein	0.9913		Pork	0.7489	Fat	0.9793	
Dough	0.8674	Fiber	0.9783		Sauce	0.7481	Sodium	0.9284	
Tomato	0.7795	Sodium	0.9549		Garlic	0.7447	Cholesterol	0.8988	
Basil	0.7569	Cholesterol	0.9449		Celery	0.7049	Fiber	0.8857	
Olive	0.7402	**Attribute**	**Value**		Pepper	0.6029	**Attribute**	**Value**	
Flour	0.5167	Bake	1.0000	Pizza	Beef	0.5564	Bake	1.0000	Chicken Wings
Pepper	0.4543	Crispy	0.9193		Carrot	0.5092	Crispy	0.9193	
Mozzarella	0.4162	Grill	0.9193		Lemon	0.5058	Grill	0.9193	
Garlic	0.360	Juicy	0.8416		Rice	0.4526	Juicy	0.8416	
		Crunchy	0.8263				Hot	0.8263	

Fig. 5 Final results showing estimated attributes, ingredients and nutrients for given food item image. (i) Pizza (ii) chicken wings

5 Conclusion

This paper attempts to make contributions in the fields of health and diet management. The purpose of this paper is to analyze the user's needs and difficulties faced by them while using the existing applications. The system can capture an image of a single food item and recognize it. It can estimate the attributes, ingredients and nutrients in real time so that users do not need to fill in the details manually. The main modules, i.e., food recognition and attribute estimation module, are implemented by using convolutional neural networks and Word2Vec tool, respectively. The supporting modules include: food logging, calorie counter and healthy recipe recommender that assist in diet management. The proposed system can be upgraded in the near future to recognize multiple food items in a single image. It can be improved to obtain the food attributes based on the portion size of the food items. More aspects such as volume and area estimation techniques can be incorporated for boosting the performance and accuracy.

References

1. Afshin, A., Sur, P. J., Fay, K. A., Cornaby, L., Ferrara, G., Salama, J. S., & Afarideh, M. (2019, May 11) Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet, 393*(10184), 1958–1972.
2. Times of India [Online]. Available https://timesofindia.indiatimes.com/india/unhealthy-diet-bigger-killer-than-tobacco/articleshow/68759153.cms.
3. Samsung Health. [Online]. Available https://health.apps.samsung.com/.
4. MyFitnessPal. [Online]. Available https://www.myfitnesspal.com/.
5. Ahmed, A., & Ozeki, T. (2015). Food image recognition by using bag-of-surf features and hog features. In *Proceedings of 3rd International Conference on Human-Agent Interaction* (pp. 179–180).
6. Hassannejad, H., Matrella, G., Ciampolini, P., De Munari, I., Mordonini, M., & Cagnoni, S. (2016). Food image recognition using very deep convolutional networks. In *Proceedings of 2nd International Workshop on Multimedia Assisted Dietary Management* (pp. 41–49).
7. Martinel, N., Foresti, G. L., & Micheloni. C. (2016). Wideslice residual networks for food recognition. [Online]. Available https://arxiv.org/abs/1612.06543.

8. Fang, S., Zhu, F., Jiang, C., Zhang, S., Boushey, C. J., & Delp, E. J. (2016, September) A comparison of food portion size estimation using geometric models and depth images, In *Proceedings of IEEE International Conference on Image Processing (ICIP)* (pp. 26–30).
9. Pouladzadeh, P., & Shirmohammadi, S. (2017). Mobile multi-food recognition using deep learning. *ACM Transactions on Multimedia Computing, Communications, and Applications, 13*(3s), 36.
10. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Proceedings of Advances in Neural Information Processing Systems* (pp. 1097–1105).
11. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. [Online]. Available https://arxiv.org/abs/1409.1556.
12. Nasi, G., Cucciniello, M., & Guerrazzi, C. (2015). The role of mobile technologies in health care processes: The case of cancer supportive care. *Journal of medical Internet research, 17*(2).
13. Marcolino, M. S., Oliveira, J. A. Q., D'Agostino, M., Ribeiro, A. L., Alkmim, M. B. M., & Novillo–Ortiz, D. (2018). The impact of mhealth interventions: Systematic review of systematic reviews. *JMIR mHealth uHealth, 6*(1).
14. Chin, C.-L., Huang, C.-C., Lin, B.-J., Wu, G.-R., Weng, T.-C., & Chen, H.-F. (2016). Smartphone-based food category and nutrition quantity recognition in food image with deep learning algorithm. In *International Conference on Fuzzy Theory and Its Applications.*
15. Jiang, L., Qiu, B., Liu, X., Huang, C., & Lin, K. (2020). DeepFood: Food image analysis and dietary assessment via deep model. *IEEE Access, 8,* 47477–47489.
16. Yunus, R., et al. (2018, December). A framework to estimate the nutritional value of food in real time using deep learning techniques. *7*, 2643–2652.
17. Meyers, A., et al. (2015, December). Im2Calories: Towards an automated mobile vision food diary. In *Proceedings of IEEE International Conference on Computer Vision* (pp. 1233–1241).
18. Ege, T., & Yanai, K. (2017). Image–based food calorie estimation using knowledge on food categories, ingredients and cooking directions. In *Proceedings of Thematic Workshops* (pp. 367–375).
19. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. [Online]. Available https://arxiv.org/abs/1301.3781.
20. Zhang, W., Zhao, D., Gong, W., Li, Z., Lu, Q., & Yang, S. (2015). *Food image recognition with convolutional neural networks* (pp. 690–693).
21. Suh, H.-J., & Lee, K-,H. (2018). Real–time calorie extraction and cuisine classification through food–image recognition. *International Journal of Grid and Distributed Computing, 11,* 69–78.
22. Temdee, P., & Uttama, S. (2017). *Food recognition on smartphone using transfer learning of convolution neural network* (pp. 132–135). Global Wireless Summit (GWS), Cape Town.
23. Islam, M. T., Karim Siddique, B. M. N., Rahman, S., & Jabid, T. (2018). Food image classification with convolutional neural network. In: *2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)* (pp. 257–262), Bangkok.
24. Harvard Health Publishing [Online] Available https://www.health.harvard.edu/blog/why-keep-a-food-diary-20190131158555.

A Deep Learning Generative Approach for Speech-to-Scene Generation

Abhijith Venugopal, Adapa Shivani, M. Neha, and H. R. Mamatha

Abstract Visualization can enhance the power of our subconscious mind. Research has proven that visualization is a very effective medium for communication since it enables humans to remember insights for a longer duration. A speech input visualizer would thus be of utmost importance because of its wide-ranging applications in areas such as Education, Engineering, Defence, Art, Game Development, Architecture, and so on. This paper introduces a real-time efficient "Speech-to-Scene" Generator using a deep learning approach. A novel combination of HTML speech recognition API and Text Conditioned Auxiliary Classifier Generative Adversarial Network has been proposed to overcome this problem. This model has been trained on a car's data set consisting of 16 distinct classes. Additionally, this model can also adapt to various other data sets. The generated images were evaluated with Inception Score and Multi-Scale Structural Similarity Index (MS-SSIM) to compare with state-of-the-art image generation technologies.

Keywords Visualization · Deep learning · Image synthesis · Audio-visual systems · Generative approach

Abhijith Venugopal and Adapa Shivan—These authors contributed equally to this work.

A. Venugopal (✉) · A. Shivani · M. Neha · H. R. Mamatha
PES University, Bangalore, India
e-mail: abhijith1998@gmail.com

A. Shivani
e-mail: a.shivani63@gmail.com

M. Neha
e-mail: nehamohan1018@gmail.com

H. R. Mamatha
e-mail: mamathahr@pes.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_59

1 Introduction

It is rightly said that "a picture can be worth a thousand words" since a photograph, painting, animation or artwork can often depict ideas faster than any other communication medium. Human's capability to visualize speech is an extremely intricate mechanism, since people's perception is usually in different ways and forms, and conveying the same to others is a herculean task. This complex process of communicating ideas can be greatly simplified through images. One scenario, where this becomes evident, is in the field of education where conveying of ideas across mediums becomes crucial. Several children find it difficult to grasp concepts through audio or text descriptions alone. Further, cases such as learning disabilities and lack of knowledge of languages worsen the ability to grasp concepts from the standard conventional method of verbal or written communication. Hence, a scene generation system would be beneficial in these fields. Moreover, traditional methods such as manual scene composition techniques are tedious and time-consuming because it requires users to learn special software or tools related to graphics.

In today's world, machine learning and deep neural networks are spreading their influence on all domains, with image generation and speech processing being no exception. In particular, a significant amount of work has been done in generating synthetic images using generative models, specifically generative adversarial network (GAN) model. GANs are used to "generate" or "create" new data. The original architecture was propositioned by Goodfellow et al. [1], wherein two models are trained in parallel, consisting of a generative model that captures the distribution of the data and a discriminative model which calculates the probability of a sample from the training data rather than the "fake" generated one. The generator tries to maximize the probability of the discriminator making a mistake.

The rest of the paper is organized as follows: Sect. 2 describes related work, Sect. 3 encapsulates the data set description, Sect. 4 gives a brief overview on the background, Sect. 5 elaborates on the implementation details, Sect. 6 describes the results achieved, Sect. 7 describes evaluation, and finally Sect. 8 describes the future work.

2 Related Work

In the recent times, there has been a significant amount of work done in the domains of scene representation, image synthesis and speech-to-text conversion. Papers such as WordsEye [2] and learning spatial knowledge for text to 3D scene generation [3] explore text to scene conversion, with emphasis on representation of spatial knowledge. This provides an insight on how positional information should be represented for efficient conversion in models. Deep learning approaches have always been favoured in the domain of image generation, and hence, there have been a large number of advancements in this field, with different concepts and models being

used. Variational autoencoders have been established to produce images similar to a sample distribution. The DRAW model [4] has encapsulated this and extended it further to produce high accuracy handwritten digits, with concepts such as "attention mechanism" to improve accuracy.

Generative adversarial networks (GANs) have always been used for a multitude of applications due to their high efficiency and accuracy in combination with diverse flexibility. Different kinds of GANs have been used in the previous researches, to synthesize images from various sources. Some examples are Style and Structure GAN [5] and Laplacian GAN (LAPGAN) [6] which use a layered architecture to refine the blurry images produced from normal GAN models. Zhang et al. [7] proposed Stack GAN model which is another proven generative model that leverages the benefit of using multiple stages, to produce highly realistic images.

Another approach which has become popular with its high performance is the concept of making the generative models more aware of their domain and data population, that is, making the model more "contextual" and "conditioned". The conditional GAN model (CGAN) [8] and auxiliary classifier GAN (AC-GAN) [9] have proven this, with the models being able to produce high-quality coherent images for different categories and classes. An extension of this concept has been used in Text Conditioned Auxiliary Classifier GAN (TAC-GAN) model [10], where text embeddings have been used as an additional criteria for generating better quality images. This approach has proven to be quite effective in text to image conversion, with highly comprehensive, coherent images being generated for text captions (Fig. 1).

3 Data Set Description

The dataset used in the implemented system is a custom car image data collection, which was specially made for image processing applications, by manually creating comprehensive and detailed captions for standard car images. With a large set of high-resolution vehicle images from Carvana Kaggle image data set [11], text summaries for the same were meticulously written, with the objective being to highlight the importance of large meaningful descriptions in image processing (Fig. 2). The images have been categorized into 16 classes, with each class having vehicles of different automobile companies, giving an overall total of around 6000 quality images. Further, each image is associated with detailed textual descriptions which highlight the features displayed in the image. Overall, the data set is a rich variety of images of different vehicle types and colours, with different orientations, positions and lighting. The detailed text captions coupled with the all-rounded images make it very useful in the task of speech-to-scene generation. These properties of the data collection have made it indispensable in this task of speech-to-scene generation, and therefore, this data set has been extensively used, to train and test the model.

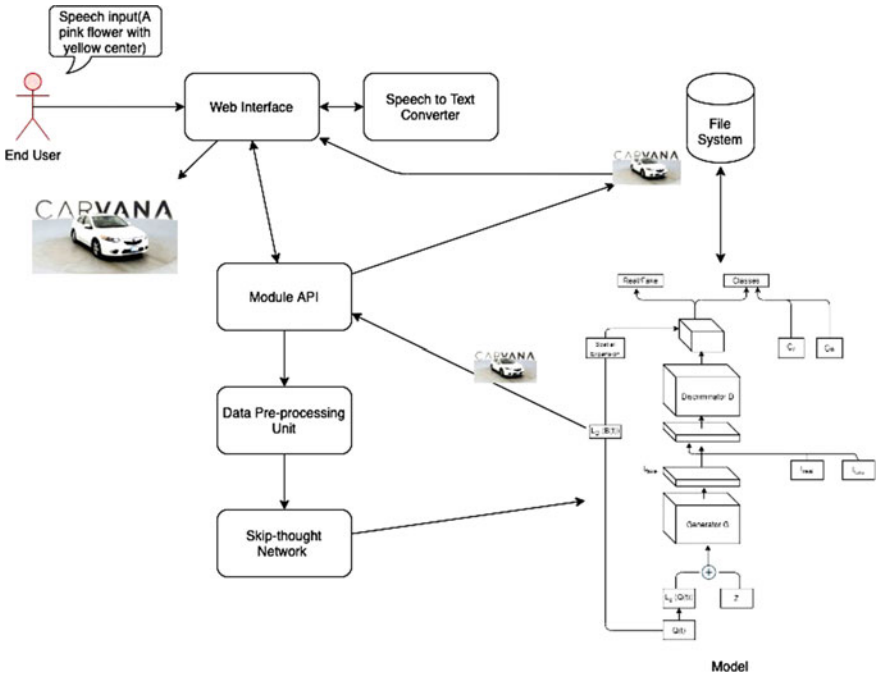
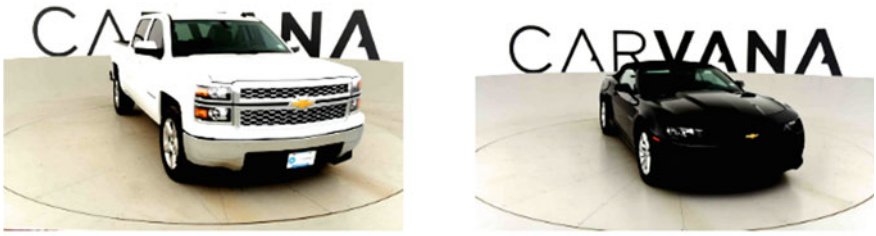


Fig. 1 Overall architecture



The body of this car is white in color.
This Chevrolet car, is of model type Silverado 1500 and is white in color
This car has white body and is medium in size.
This car has a body of white color and it is a sedan.
This car has white wing mirror cover.
This car has black tyre with silver rim.

The body of this car is black in color.
This Chevrolet car, is of model type Camaro and is black in color.
This car has black body and is medium in size
This car has a body of black color and it is a sedan
This car has black wing mirror cover.
This car has black tyre with silver rim

Fig. 2 Dataset samples. The images of resolution have been obtained from Kaggle Carvana dataset [11], for which textual descriptions have been developed

4 Background

This section gives information about generative models in general, with detailed description of TAC-GAN architecture in specific.

4.1 Generative Adversarial Networks (GANs)

Generative adversarial networks (GANs) are deep learning generative models which, as the name suggests, are primarily used to "generate" or "create" new data. They are essentially algorithmic architectures consisting of a couple of neural networks, which are forced to compete against one another, eventually producing a model that can generate new artificial data which is identical to "real data".

A GAN consists of a generator network which creates new synthetic data (generally data from some known distribution with noise factor) and a discriminator network responsible for determining if the data passed to it is "real" or "fake". These networks are put in a standoff situation, where the discriminator tries to accurately find the false data, whereas the generator tries its best to deceive the discriminator. The essential loss function used to train the simple GAN model is called Minimax loss given by:

$$Ex[\log(D(x))] + Ez[\log(1 - D(G(z)))] \tag{1}$$

where in:

- $D(x)$ is the discriminator's estimation of real data x being real.
- Ex is the expected value over all real data.
- $G(z)$ is the generator's output.
- $D(G(z))$ is the discriminator's estimation of fake data being real.
- Ez is the expected value over all random inputs to the generator.

The generator tries to minimize this function while the discriminator tries to maximize it. Convergence occurs at Nash equilibrium in most cases, beyond which further improvement is not possible.

4.2 Text Conditioned Auxiliary Classifier Generative Adversarial Network(TAC-GAN)

TAC-GAN is a text to image generative model used to generate images from their textual descriptions. Whereas an ordinary GAN discriminator is given the task of only finding out if an image is real or fake, the discriminator here also associates a textual description with the image it gets. Therefore, it learns to identify and associate textual descriptions with their corresponding images, which is passed to the discriminator

via the loss function(since the generator has to produce better images to reduce loss and fool the discriminator). Therefore, in this architecture, each image has a corresponding vector of textual descriptions which is used as an additional factor in identifying its class.

Technically, TAC-GAN is a variation of auxiliary classifier generative networks in which every generated image is closely linked with a class label. In an AC-GAN, every produced images have a class "C" along with a noise vector "z", which are used by the generator in image synthesis. TAC-GAN extends this concept one step further, with application of text embedding in conjunction with noise vectors to facilitate efficient image synthesis. Therefore, in this architecture, every data instance is a tuple (I, Q), where "I" is the image and "Q" is the latent representation of the text description set associated with the image.

5 Implementation Details

This section will thoroughly describe various components of the proposed model, with summaries of technologies and architecture used to achieve it. Figure 1 depicts the high-level overall architecture of the model for better understanding.

5.1 Speech-to-Text Conversion

Speech-to-text conversion refers to the efficient and legitimate conversion of speech to text. Human speech processing is a complex and intricate domain, with research being done constantly on how to map speech patterns in an effective manner. Being a core component of the proposed speech-to-scene generator, efficient conversion of user speech input is of vital importance to the accurate functioning of the system. Hence, different technologies and tools were looked into, with focus being on accuracy of speech-to-text conversion, ease of use, availability, cost, ease of implementation and flexibility, and so on. Considering a combination of these factors, HTML5 Speech Recognition API was chosen as speech-to-text conversion method.

5.2 HTML5 Speech Recognition API

A HTML5 Speech Recognition API is a flexible and convenient API which can be easily integrated with JavaScript code. In simple terms, it enables JavaScript to access a browser's audio stream and convert the speech sequence to its corresponding text transcript. Internally, this module uses Google's speech recognition engine for its processing. This API is extremely simple to use and overcomes most of the shortcomings of its counterpart speech-to-text technologies such as Julius, Wav2Letter++

and DeepSpeech. A highlight, which helps in flexibility and extension of the system, is that it supports more than 61 languages including various Indian languages such as Telugu, Kannada, Tamil, Marathi, Malayalam, Hindi and Gujarati. Moreover, it supports various English dialects, thereby improving chances of efficient detection of speech inputs. Each resultant string and its alternative have a confidence value which indicates how confident the speech recognition engine is of the result. The system uses these features of the model effectively, by encompassing the module in the Website front end, thereby enabling quick speech-to-text processing, simulating real-time conversion.

5.3 Text-to-Image Conversion

This segment involves the second phase of the system, to convert the result of speech processing into a comprehensive, meaningful image. For this purpose, a deep learning model was developed, using concept of TAC-GAN and skip-thought vectors, which functions as an Application Programming Interface back-end, enabling flexibility and re-usability of the system.

5.3.1 Skip-Thought Vectors

Converting textual information to mathematical vectors and representations is an important step in machine learning. Skip-thought vectors were introduced in 2015 to provide sentence-level vectors that are learned in an unsupervised manner. Skip-thought vectors, or skip thoughts, provide fixed-length text representation which enables us to replace any sentence with its corresponding vector of numbers. Skip thoughts are an extension of the skip-gram model, incorporating the same ideology as that of skip grams, but for sentences. Instead of predicting context by surrounding words, skip thoughts try to predict target sentences using surrounding sentences.

Skip-thought vectors are generated using an encoder-decoder framework, where there are two decoders for the previous and next sentences and the encoder is used to form a representation for the current sentence. Therefore, the idea is that the representation provided by the encoder for the current sentence should encapsulate semantic meaning to accurately predict the previous and next sentences in context.

5.3.2 Deep Learning Network

Using the concept of passing text embedding along with images to generative model as depicted by TAC-GAN architecture, a deep neural network was designed to implement a generative model comprising of a convolutional generator and discriminator pair, which accepts data as a tuple consisting of image and its embedding. The loss functions were tweaked in a manner to couple the two attributes, so that the image and

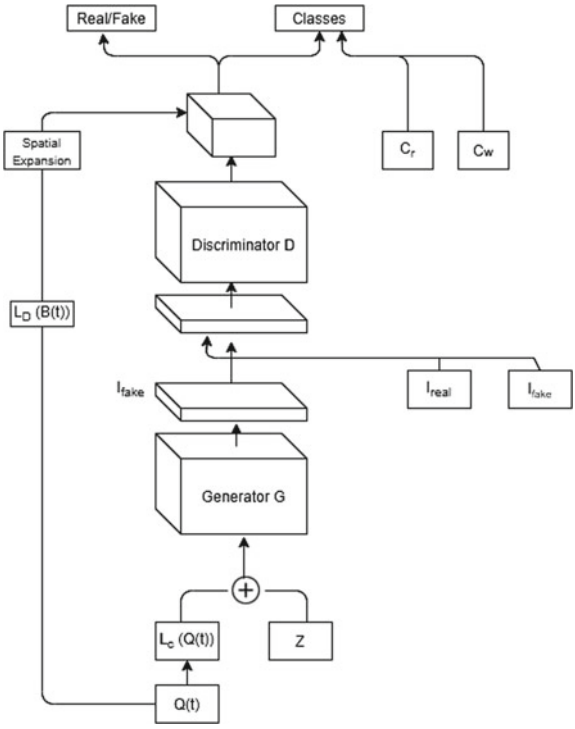


Fig. 3 Proposed model architecture. Here, z is the noise vector, I_{real} and I_{fake} refer to real and fake images, where I_{fake} is produced by the Generator model G. Q_t is the text embedding of caption t. C_r and C_w are class labels (one hot encoded) of I_{real} and I_{fake}. L_G and L_D are the two deep neural networks that produce latent representation for text embedding Q_t

text embedding would be strongly associated, and this information would be passed on to the generator. This enables the generator to get acclimatized to the embedding associated with the image, with the end goal of being able to re-create images with just the embeddings. Figure 3 depicts the detailed proposed model architecture. The specifications of the deep learning model developed are as follows:

- Network Architecture: The generator network consists of three transposed convolutional neural layers. The number of filter maps for these layers are 256, 128 and 64, respectively. Each layer's output size is double to that of its input image size. The last layers output is fed as input to the discriminator. The discriminator consists of three deconvolutional neural layers, of 128, 256 and 384 filter maps, respectively. The kernel size for all layers is 5×5 until the second last layer, Md. The output of Md is concatenated with the text embedding(which is made linear and then tiled), lr, and this is then fed to the last layer consisting of 512 filter maps having size 1×1 and stride 1.
- Loss function: Discriminator Loss:

$$Lds = H(Ds(Ir, lr), 1) + H(Ds(If, lr), 0) + H(Ds(Iw, lr), 0) \qquad (2)$$

where *Lds* denotes the training loss related to the source of the input, i.e. real, fake or wrong. This is obtained as the sum of binary cross entropy, denoted by *H*, between the discriminator's output and the desired value for each of the images.

$$Ldy = H(Dy(Ir, lr), Qr) + H(Ds(If, lr), Qf) + H(Ds(Iw, lr), Qw) \quad (3)$$

where *Ldy* denotes the training loss of the discriminator on classifying the images with respect to a dataset *y*, which contains an embedding *Q* for each of the images in it. *Q* is the vector which contains relevant information, such as position of object, attributes intensity and so on. The discriminator tries to minimize

$$Lds + Ldy \quad (4)$$

Generator Loss: The generator loss is given by

$$Lgs + Lgy \quad (5)$$

where:

$$Lgs = H(Ds(If, lr), 1) \quad (6)$$

$$Lgy = H(Dy(if, lr), Qf) \quad (7)$$

Here, *Lgs* takes into account the output of the discriminator for the generated image, while *Lgy* considers the expected embedding *Qf* for the generated image.

6 Results

Some of the images that were generated after training the model for 108 epochs are depicted in Figs. 4 and 5.

Fig. 4 This car has grey body and is medium in size

Fig. 5 This car has black body and is medium in size

7 Evaluation

Evaluating GAN generator models has been a challenge since its conception, because the GAN generator model is trained with another model called as a discriminator, which finds about the authenticity of the generated image. The generator and discriminator compete with one another, due to which the loss measurements done for each of them keep fluctuating, in a manner similar to a zero sum game. Due to this, analysing the improvements in a quantitative method becomes difficult and no direct measurement of loss gives sufficient information on the efficiency of the model. In the initial stages of training, the model was saved every few epochs and synthetic images were generated to validate the improvements in the model. As the model made progress and learnt varied features, more comprehensive approaches were used to evaluate the model. The testing methodologies that were used to evaluate the model are Inception Score and Multi-Scale Structural Similarity Index (MS-SSIM). Inception Score has been used to calculate the ability of the proposed model to discriminate the generated images, whereas MS-SSIM was used to evaluate the heterogeneity of the produced images. In comparison with state-of-the-art TAC-GAN model, the mean Inception Score of the proposed model after 89 epochs, as shown in Table 1, was 3.19 with a standard deviation of 0.0979. The results depicted here are for 89 epochs, with better MS-SSIM score expected on further training aided with greater computation power. Thus, the results indicate that the proposed model is suitable for accurate image visualizations.

Table 1 Inception score comparison with state-of-the-art models

S. No.	Model	Inception score
1	Proposed model	3.139
2	TAC-GAN	3.45 ± 0.05
3	StackGAN	3.20 ± 0.01
4	GAN-INT-CLS	2.66 ± 0.03

8 Future Work

The proposed model is a highly flexible and extensible one. In addition to the work done, the model can be extended to add a feedback mechanism, wherein the user can provide ratings to improve the model's performance. Besides these features, exploration can be done concerning the model itself, wherein deep learning approaches such as attention mechanism and capsule networks can be used for better performance. The model is currently generating 2D images for speech inputs, and another add-on would be to extend it to the next dimension, wherein the model would be able to generate 3D models to capture information specified to it.

References

1. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. ... Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672–2680).
2. Coyne, B., & Sproat, R. (2001, August). WordsEye: An automatic text-to-scene conversion system. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 487–496).
3. Chang, A., Savva, M., & Manning, C. D. (2014, October). Learning spatial knowledge for text to 3D scene generation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 2028–2038).
4. Gregor, K., Danihelka, I., Graves, A., Rezende, D. J., & Wierstra, D. (2015). *Draw: A Recurrent Neural Network for Image Deneration*. arXiv preprint arXiv:1502.04623.
5. Wang, X., & Gupta, A. (2016, October). Generative image modeling using style and structure adversarial networks. In *European Conference on Computer Vision* (pp. 318–335). Cham: Springer.
6. Denton, E. L., Chintala, S., & Fergus, R. (2015). Deep generative image models using a laplacian pyramid of adversarial networks. In *Advances in neural information processing systems* (pp. 1486–1494).
7. Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., & Metaxas, D. N. (2017). Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 5907–5915).
8. Mirza, M., & Osindero, S. (2014). *Conditional generative adversarial nets*. arXiv preprint arXiv:1411.1784.
9. Odena, A., Olah, C., & Shlens, J. (2017, August). Conditional image synthesis with auxiliary classifier gans. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70* (pp. 2642–2651). JMLR. org.
10. Dash, A., Gamboa, J. C. B., Ahmed, S., Liwicki, M., & Afzal, M. Z. (2017). *Tac-gan-text conditioned auxiliary classifier generative adversarial network*. arXiv preprint arXiv:1703.06412.
11. Carvana Image Masking Challenge. https://www.kaggle.com/c/carvana-image-masking-challenge/data. Accessed 14 2020

A Comparative Analysis of Genes Responsible for Cataract

Samar Jyoti Saikia and S. R. Nirmala

Abstract A gene is the basic unit of genetic material. It is made up of a sequence (or piece) of DNA. Each gene controls a particular feature or has a particular function in the body. There are various eye diseases which are the result of genetic inheritance and which lead to blindness in infants, children, and adults. Out of every three blind people in the world, one is from India. 7.4 million, i.e., 62% have cataract-related blindness in India. This work focuses on developing a mathematical model to analyze genes responsible for various types of cataract. Another perspective of this work is to establish the relationship of genetic factors with other factors of the diseases. In this work, a comparative study is done among different cataract samples following intragenic and intergenic approach.

Keywords Cataract · HSF4 gene · ABHD12 gene · Nucleotide density · Codon density and amino acids

1 Introduction

There has been a virtual explosion of genomic sequence data analysis with numerous genomes in various stages of sequencing and annotation in the last few years. Human genome project (HGP) has been a globally known framework in which researchers across the world collaborated to determine the whole genetic information in the human body. It further explored different aspects of the evolution of the human genome and the revolution that has occurred in the human life and medical sciences in many ways. All the human chromosomes have been sequenced after the completion of human genome project in the year 2003 [1]. To know more about biology and the evolutionary relationships between these genomes, accurate annotation of these

S. J. Saikia (✉) · S. R. Nirmala
Gauhati University, Guwahati, Assam 781014, India
e-mail: samar.saikia@dbuniversity.ac.in

S. R. Nirmala
e-mail: nirmalasr3@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_60

genomes is essential [2]. Thus, the development of reliable automated techniques for partitioning them into genes, promoters, regulatory elements, and intergenic region became imperative.

There are various eye diseases which are the result of genetic inheritance and which lead to blindness in infants, children, and adults. Identification of genes related to eye diseases is yet not completed. To predict and detect the presence of eye disease-related genes is an important bioinformatics task. Soft computational tools play a vital role in the prediction of those genes. Congenital cataracts, glaucoma, retinal degeneration, eye malformations, retinitis pigmentosa, etc., are some inherited eye diseases. The adaptive and learning capabilities of soft computational tools make the gene identification process easier, less time consuming, and inexpensive.

2 Background

Proteins are the fundamental molecules of all organisms. During the reproduction, protein building instructions are largely controlled by an organic molecule called nucleotides. By joining nucleotides, nucleic acids are formed. Deoxyribonucleic acids (DNA) and ribonucleic acid (RNA) are the two types of nucleic acids. DNAs are the most important biochemical molecule of the human body. After the invention of the microscope, it has been known that body of an organism is made up of cells. The cell of an organism consists of a small piece of information holding matter that contains the blueprint of formation of cell inside the body. In 1944, this small encyclopedia of the cell was discovered as a chemical compound which is denoted as deoxyribonucleic acid (DNA). DNAs can be regarded as the instruction manual to build a life. DNA is a molecule that carries hereditary information from one generation to another. The DNA present in the cell nucleus is known as nuclear DNA and the DNA found in the mitochondria is known as mitochondrial DNA [3]. DNA molecule is a combination of a deoxyribose sugar, phosphate, and nitrogenous bases. DNA stores information as a code in terms of chemical bases, namely adenine (A), guanine (G), cytosine (C), and thymine (T) [4]. The DNA molecule of higher organisms consists of two strands. These strands are linear arrangement of some similar units known as nucleotides. Adenine (A), thymine (T), guanine (G), and cytosine (C) are four different bases that comprise of DNA molecule. By forming base pairs, the two strands are held together. There is a fixed set of rules for base pairing. The two rules are—adenine always pairs with thymine (A-T) and guanine always pairs with cytosine (G-C). The total number of base pair is known as genome size. Genes are the carriers of genetic information in an organism. A gene is a segment of DNA.

3 Cataract

Clouding of the lens those results in a blurred image is referred to as a cataract. The light can be passed through the lens which is mainly composed of water and proteins [5].

Cataract can be categorized into three types:

(i) Nuclear cataract occurs in the center of the lens
(ii) Cortical cataracts, this type of cataract starts in the cortex, i.e., in the periphery of the lens and this gradually extend to the center of the lens
(iii) Sub-capsular cataracts are those in which the opacities are concentrated beneath or within the capsule of the lens.

Mutation in protein which is present in the eye lens can cause cataract which is known as congenital cataract [6]. In case of congenital cataract, there is no progression of lens opacities which can lead to visual impairment. Different mutations in the same gene or same mutation in different gene may lead to congenital cataract. The genes responsible for cataract are shown in Table 1.

From the above-mentioned genes in Table 1, in the proposed work, we have considered HSF4 and ABHD12 genes. The full form of HSF4 is heat shock transcription factor. The expression heat-shock proteins such as lens αB-crystallin are activated by this gene [7]. Mutations in HSF4 can lead to autosomal dominant and recessive cataracts. The other name of dominant cataract is lamellar which is seen in first few years of birth [8, 9], where as the congenital onset is responsible for recessive cataracts [10].

The Alpha/Beta Hydrolase Domain Containing 12 (ABHD12) is a protein coding type of gene. Mutation in this gene can lead to polyneuropathy, hearing loss, ataxia, retinitis pigmentosa, cataract (PHARC) which is an inborn error of endocannabinoid metabolism.

Table 1 Genes responsible for cataract

Sl. No.	Gene	Disease
1	HSF4	Autosomal dominant or recessive cataract
2	ABHD12	Responsible for polyneuropathy, hearing loss, ataxia, retinitis pigmentosa, cataract (PHARC)
3	PITX3	Posterior polar cataract associated with anterior segment mesenchymal dysgenesis (ASMD) including corneal opacity, iris adhesions, and optic nerve abnormalities
4	FOXE3	Cataract as part of an (ASMD) spectrum
5	EYA1	Cataract with dysgenesis of the anterior segment of the eye
6	CHX10	Cataracts associated with microphthalmia and iris defect
7	MAF	Cataracts associated with ASD
8	PAX6	Anterior segment malformations including cataract

4 Methodology

4.1 Analysis of Genes

In the proposed work, HSF4 and ABHD12 and samples from two patient genes have been considered for the genetic analysis. These two genes are primarily responsible for different types of cataract. HSF4 is responsible for autosomal and dominant cataract and ABHD12 is responsible for PHARC. In this work, we have analyzed genes using two approaches–intragenic and intergenic. In the intragenic approach, mutation within the same gene is considered, whereas, in the intergenic approach, mutation in the different genes has been considered, i.e., in the intragenic approach, samples of the only HSF4 gene are compared, whereas, in the intergenic approach, samples of HSF4 gene and ABHD12 gene are considered. Figure 1 represents the block diagram of the proposed model. Following samples are considered during the analysis:

(1) Nucleotide sequence of healthy HSF4 gene (Sample 1)
(2) Sample 2 from patient (mutated form of HSF4)
(3) Sample 3 from patient (mutated form of HSF4)
(4) Nucleotide sequence of healthy ABHD12 gene (Sample 4).

The algorithm to analyze the samples:

1. Sample collection
2. Generation of coding scheme
3. Similarity matching
4. Use of statistical methods.

1. **Sample Collection**. Table 2 represents the samples considered for the work. The gene samples are collected from National Center for Biological Information (NCBI) database.

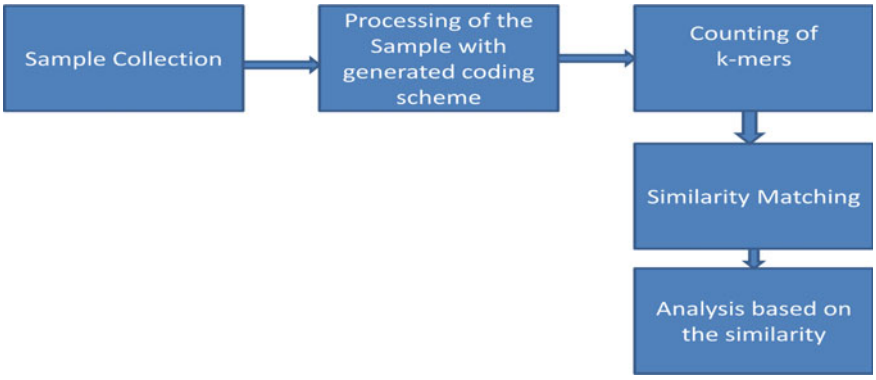


Fig. 1 Graphical representation of the proposed model

Table 2 Samples considered in the proposed work

Sl. No.	Sample	Accession No.
1	Sample 1(healthy HSF4)	NC_000016
2	Sample 2 (patient Sample1)	NM_001040667
3	Sample 3 (patient Sample2)	NM_001538
4	Sample 4 (healthy ABHD12)	NG_028119

2. **Generation of Coding Scheme**. A coding scheme is generated to convert the samples into a mathematical format. These codes are generated based on the molecular weight, solubility and the melting point of adenine, thymine, guanine, and cytosine. The code generated from the coding scheme is as follows:
   Adenine (A): 8
   Guanine (G): 3
   Thymine (T): 2
   Cytosine (C): 1
3. **Similarity matching**. Similarity matching is done in two ways—intragenic and intergenic. Algorithm for Similarity Matching is

   - *Counting of k-mers*. Certain nucleotide strings often appear surprisingly often in small regions of the genome. This is often because certain proteins can only bind to DNA if a specific string of nucleotides is present, and if there are more occurrences of the string, then it is more likely that binding will successfully occur. It is also less likely that a mutation will disrupt the binding process. The term k-mer is used to refer to a string of length k.
   - The mathematical formula is used to calculate the *k*-mers—

$$Z = p^k \tag{1}$$

     where Z = number of k mer, p = number of input, and k = value of mers, 1 to n. In our work, we have considered $p = 4$ because there are four nucleotide bases such as A, T, G, C, k value is considered till three because the combination of three nucleotide bases can lead to amino acids which can be used for further analysis and Z is dependent on k.
   - Matching of the number of k-mers for forward and reverse complement: Density of *k*-mers is considered by comparing the number of *k*-mers of each sample.
   - Detection of mutation among the samples: Mutation is detected by considering the amino acid sequence. Nucleotide sequences of the samples are considered to amino acids by considering triplet of nucleotides. The converted amino acid sequence is further used for statistical analysis.

5 Experimental Results

5.1 Calculation of Nucleotide Density

Not all portions of the DNA sequences are coding. The template for a protein is represented by the coding zones. In the DNA sequence of human, only 3–5% is coding, i.e., they form the gene. Sections of a DNA sequence with a high percent of $A + T$ nucleotides usually indicate intergenic parts of the sequence, while low $A + T$ and higher $G + C$ nucleotide percentages indicate possible genes. The bar diagram shows the nucleotide density in Fig. 2. From Fig. 2, we can draw the following observation:

- Nucleotide density of HSF4 healthy sequence is more than that of Sample 1 and Sample 2.
- The density of all the four nucleotides in HSF4 healthy gene is in the range of above 1000.
- Nucleotide density of Sample 1 and Sample 2 falls below 1000.

Figure 3 shows the nucleotide density of ABHD12 gene (Sample 4) versus patient sample (Sample 2 and Sample 3). Here also, we can see that the nucleotide density is more in the normal gene than that of patient samples

Then, we have compared the codon density or trimer density among the different samples. Figures 4 and 5 show the codon density of Sample 1 (HSF4 healthy gene) versus Sample 2 (patient Sample 1), and codon density of Sample 1 (HSF4 healthy

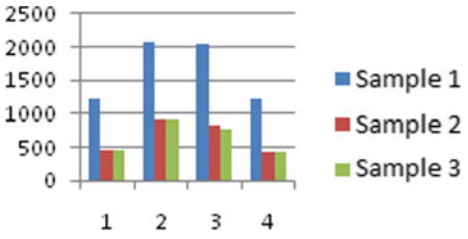


Fig. 2 Nucleotide density of HSF4 healthy sequence (Sample 1) versus patient sample (Sample 2 and Sample 3)

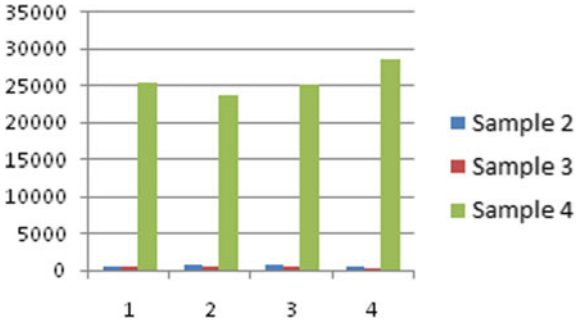


Fig. 3 Nucleotide density of ABHD12 gene (Sample 4) versus patient sample (Sample 2 and Sample 3)

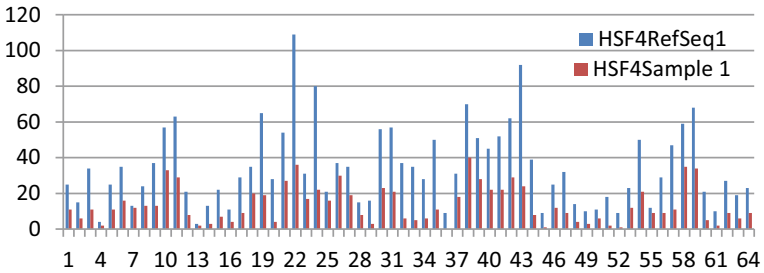


Fig. 4 Codon density of HSF4 gene (Sample 1) versus patient Sample 1 (Sample 2)

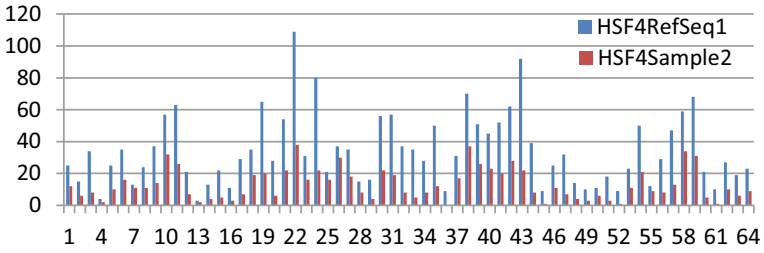


Fig. 5 Codon density of HSF4 gene (Sample 1) versus patient Sample 2 (Sample 3)

gene) versus Sample 3 (patient Sample 2) when $k = 3$. We can conclude from Figs. 3 and 4 that number of all the trimers are more in healthy HSF4 gene than that of patient samples and trimer GAT is absent in patient samples.

5.2 Amino Acid Conversion

The protein-coding sequence from the gene sequence has been extracted and converted it to the amino acid sequence for the protein. The relative amino acid composition of a protein gives the characteristic profile of the protein. The protein profile can be used to identify a protein using the amino acid composition, atomic composition, and molecular weight.

Figures 6 and 7 show the amino acid concentration of Sample 1 (HSF4 healthy gene) versus Sample 2 (patient Sample 1), and amino acid concentration of Sample 1 (HSF4 healthy gene) versus Sample 3 (patient Sample 2). From the figures, we can conclude that numbers of all amino acids are more in HSF4 healthy gene than that of patient samples.

Next, we compared the nucleotide density, codon density, and amino acid concentration between the two patient samples. Figures 8, 9 and 10 show the comparison between patient Sample 1 and patient Sample 2 in terms of nucleotide density, codon

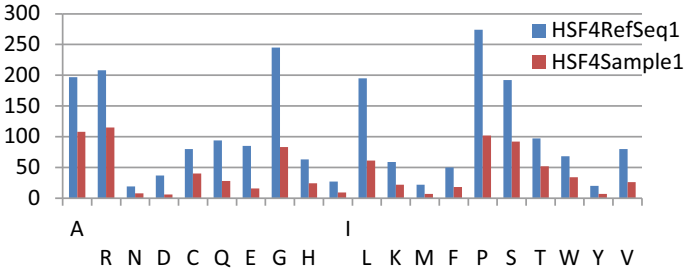


Fig. 6 Amino acid representation of Sample 1 (healthy HSF4) versus Sample 2 (patient Sample1)

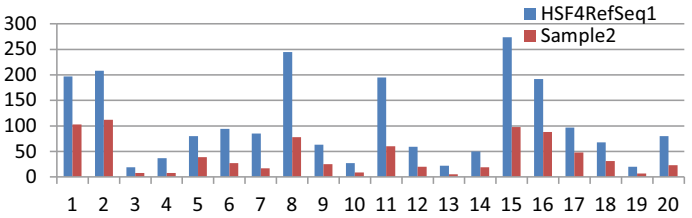


Fig. 7 Amino acid representation of Sample 1 (healthy HSF4) versus Sample 3 (patient Sample 2)

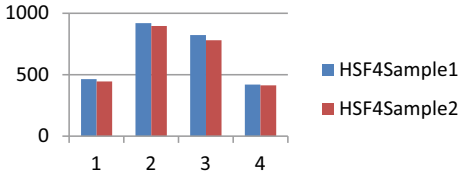


Fig. 8 Nucleotide density of patient Sample 1 versus patient Sample 2

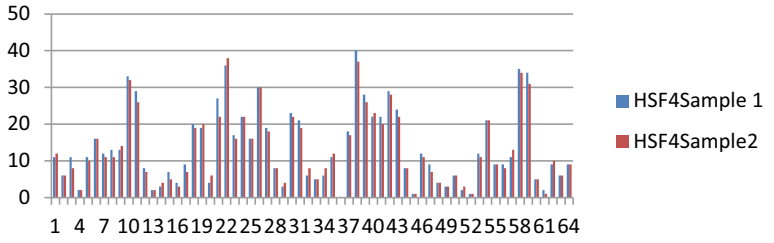


Fig. 9 Codon density of patient Sample 1 versus patient Sample 2

density, and amino acid concentration. From comparison, we can conclude that the concentration of nucleotides, codons, and amino acid is almost same in both the patient samples. Trimer GAT is absent in both the samples

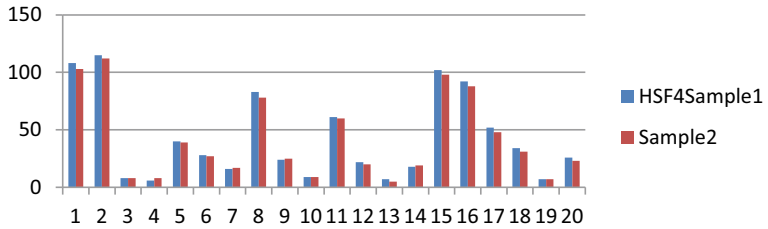


Fig. 10 Amino acid concentration of patient Sample 1 versus patient Sample 2

6 Conclusion

A comparative study is done among the samples by following intragenic and intergenic approach. When healthy gene is compared with the patient samples it shows mutation in more than 100 position which leads to cataract. When patient samples are compared it shows similarity between the samples which implies that due to some specific mutation at specific position in HSF4 and ABHD12 genes leads to cataract. In case of intergenic approach, two different genes are considered which shows huge amount of dissimilarity between them.

One of the important goals of bioinformatics is to predict the function of unknown and unclassified genes. One of the directions involves identification of the nearest classified genes using different data sources, such as protein–protein interaction data, microarray gene expressions and protein sequences. In eye-related diseases, genetic factors play a very important role, causing of blindness among infants, children, and adults. The linkage between genome structure and eye-related diseases can be explored using adaptive and learning capabilities of soft computing tools and that can help us in the diagnostic mechanism.

References

1. Bandyopadhyay, S., Maulik, U., & Roy, D. (2008). Gene Identification: Classical and computational intelligence approaches. *IEEE Transactions on Systems, Man and CyberneticsC, 38*, 55–68.
2. McElwain, M. (2007). A critical review of gene prediction software. *BioClinica, 218,* 110.
3. França, L. T. C., Carrilho, E., & Kist, T. B. L. (2002). A review of DNA sequencing techniques. *Quarterly Reviews of Biophysics, 35*(2), 169–200, Cambridge University Press, https://doi.org/10.1017/S0033583502003797.
4. Pettersson, E., Lundeberg, J., & Ahmadian, A. (2009). Generations of sequencing technologies, *Genomics, 93*(2), 105–111.
5. Eye diseases, risk factors for eye diseases and injury: Literature Review, December 2008 or https://nationaldrugstrategy.gov.au/internet/publications/publishing.nsf/Content/ageing-eyehealth-risk-factors.htm~ageing-eyehealth-risk-factors-4.htm~ageing-eyehealth-risk-factors-4-3.htm.
6. Hejtmancik, J. F., & Smaoui, N. (2003). Molecular genetics of cataract. In B. Wissinger, S. Kohl, U. Langenbeck (Eds.), *Genetics in ophthalmology* (pp. 67–82). S. Karger; Basel.

7. Somasundaram, T., & Bhat, S. P. (2004). Developmentally dictated expression of heat shock factors: Exclusive expression of HSF4 in the postnatal lens and its specific interaction with alphaB-crystallin heat shock promoter. *Journal of Biological Chemistry, 279,* 44497–44503.
8. Bu, L., Jin, Y. P., Shi, Y. F., et al. (2002). Mutant DNA-binding domain of HSF4 is associated with autosomal dominant lamellar and Marner cataract. *Nature Genetics, 31,* 276–278.
9. Eiberg, H., Marner, E., Rosenberg, T., & Mohr, J. (1988). Marner's cataract (CAM) assigned to chromosome 16: Linkage to haptoglobin. *Clinical Genetics, 34,* 272–275.
10. Forshew, T., Johnson, C. A., Khaliq, S., et al. (2005). Locus heterogeneity in autosomal recessive congenital cataracts: Linkage to 9q and germline HSF4mutations. *Human Genetics, 117,* 452–459.

A Framework for Banana Plantation Growth Using Blockchain Technology

B. Geethanjali and B. L. Muralidhara

Abstract Bananas are stacked with a rich source of nutrients and minerals that drive it to be one of the major staple crop consumption in India. The integrity and quality of bananas during its growth are a major concern for society. All the stages of the growth of the banana fruit are vital. Protection of the fruit starting from the planting stage until the marketing stage is crucial, since, the quality of banana fruit depends on how well it is supervised, during all the stages of its growth. As a solution, we propose a Banana Agro System, and the system monitors the growth of the banana fruit and its supply chain. Blockchain technology is used to store the attributes of all the growth stages of banana. Blockchain provides a secure method of storing and overseeing the data, which encourages the development and utilization of data-driven transformation for the smart farming systems.

Keywords Blockchain · Banana · Block · Farmer · Participant

1 Introduction

In recent times, blockchain is one of the most talked-about technology. Bitcoin is purely the first mainstream reflection of its potential [1]. Looking back to the second half-century of Internet technologies and architectures, one might observe a trend of fluctuation between the centralization and decentralization of computing power, storage, infrastructure, protocols, and code. Currently, we are witnessing the transition from centralized computing, storage, and processing to decentralized architectures and systems. Distributed ledger automation is one of the key improvements making this shift possible. A disbursed ledger is a kind of fact structure that resides across multiple pc devices, normally spread throughout places or regions.

B. Geethanjali (✉) · B. L. Muralidhara
Department of Computer Science and Applications, Bangalore University, Bangalore, India
e-mail: geethanjalibkrish@gmail.com

B. L. Muralidhara
e-mail: murali@bub.ernet.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_61

A blockchain is a peer-to-peer distributed ledger, formed by consent, combined with a system for sensible contracts. Blockchain is used to build a present-day generation of transactional applications that establish trust, liability, and transparency at their core. Blockchain exists as a specific kind or set of distributed ledger technologies that constructs a chronological chain of blocks, therefore the name "blockchain". A block refers to a group of transactions that are bundled along and value-added to the chain at the same time. It is primarily a public written account of who owns what and who transacts what. The transactions get fasten through cryptography, and over a while, transaction record history gets locked in blocks of information that are then cryptographically connected and secured. This creates an immutable, unforgeable record of all of the transactions throughout the network.

Nowadays, businesses of all types are becoming inventive with the blockchain technology, as it is often used to track, record, and verify trades of nearly anything that holds value. From e-commerce to cloud storage to choose, companies in all industries are starting to see blockchain's potential and integrating this technology into their functioning. Blockchain can be adopted in several logistics network chain management systems, to form a decentralized network where it will offer transparency, security, neutrality, and reliableness of all the operations being done in a logistics network system chain [2, 3]. Almeida contributes to the fact that 60% of the blockchain research topic addresses the food supply chains [4]. Tian contributes to the information of the agri-food supply chain, which establishes a healthy market environment based on RFID and blockchain technology. It covers the process of gathering data of agri-food supply chain, real-time tracking of the agri-food products, and traceability management for the quality and safety of the agri-food [5].

One of the important fruits consumed worldwide is banana, banana is a popular fruit in view of the fact that it is highly nutritive fruit, and almost all the parts of banana are consumed and useful in many ways. Such as the delicate stem of banana, which bears the inflorescence, is extricated by harvesting the leaf sheath of the pseudostem and utilized as a vegetable, the banana fiber is used to produce things like sacks, pots, wall hangers, rope, and paper, and the banana leaves are used as sterile eating plants. The fruit also has great traditional and religious values [6]. In consequence with the significant utilization of the banana fruit, the quality of the fruit is dominant. To keep track of the quality of banana in all of its growth stages, we have proposed a Banana Agro System. This system provides an overall scheme of banana growth and its supply chain system based on the blockchain technology to assure the quality of the banana. The system helps in increasing the confidence of both farmer and consumer.

2 Banana Agro System

Banana is a staple food, which has to be consumed within a few days. Even though farmers cultivate healthy bananas, farmer's do not store any data regarding the method of cultivation, the growth stages, the type and quantity of manures and pesticides used,

the date of sowing, harvesting, storage and marketing, the information on any disease attacked to banana during its cultivation, and the method used to ripen banana. Since all these details are not stored and cannot be produced when in need, farmers cannot discern the quality of their produce. And thereby farmers fail to reach the right market and get a fair price for their produce.

We propose a Banana Agro Model as a solution. The model uses blockchain to store every detail of banana growth stages. The system is designed in three layers: data sensing layer, data storage layer, and application layer.

Data Sensing layer: A range of different kinds of sensors such as temperature, humidity, soil moisture sensors are used to collect real-time agricultural land environmental data. The sensors are set up in the farmland and monitored. The sensors are read three times a day throughout the growing season of banana and are stored on microprocessors. At the onset of the harvesting stage, the mean of each of the temperature, humidity, and soil moisture data is calculated and stored on the blockchain.

Data Storage Layer: This is a principal layer in the system. Blockchain is used for decentralized storage of data. A first block called genesis block is created, which consists of index value as 0, previous hash value as series of 0's, as it is the first block, hash value, a nonce value that is generated based on hash value and a random number, and a transaction array (Fig. 1).

After the generation of the first block, further data about the growth stages of banana, farmland, and farmer is added as blocks using the transaction array. The farm license number is used as a vital key in adding data to the blocks. The farm license number is hashed using the sha-256 hashing algorithm. Hashing the farm license number is important in view of the fact that, blockchain is a decentralized storage system and we need to solitary the information only for the peer participants of blockchain. The peer participants of blockchain are farmers, wholesale dealers, retail dealers, and consumers.

The data is stored in the couch database in the blockchain. The kind of data stored is illustrated below in json format:


```
{
    "blocks": [
        {
            "index": 0,
            "previousHash": "0000000000000000",
            "hash": "00002818703517bab21046d807a3fc0284b8a05979c
e48baa40ed2eeeadd3b92",
            "nonce": 4190,
            "transactions": []
        },
```


Fig. 1 Genesis block

{
“farmLicenseNumber” : “farm001”,
“farmerName” : “Rangappa”,
“farmGeolocation” : “ 12.036756,76.7875432”,
“farmDimension” : “10 Acres”,
“soilType” : “black soil”,
“bananaSownDate” : “03/08/2019”,
“typeOfBanana” : “elakki”,
“avgTemperature” : “26 degree celcius”,
“avgHumidity” : “54%”,
“avgSoilMoisture” : “65%”,
“manurePercentUsed” : “10%”,
“pesticidePercentUsed” : “12%”,
“harvestDate” : “05/03/2020”,
“storageType” : “cold-storage”,
“numberOfDaysStored” : “12”,
“readyToSellDate” : “17/03/2020”}

All these data are not stored at a time in the block. Farm license number, farmer name, farm geolocation, farm area, soil type, type of banana, and banana seed/sapling sown date are stored first. Geolocation is an estimation of the geographic location of an area. Geolocation is essential for precise field navigation. Each farmland can be set up with a geofence. The geolocation provides the location of the farm in terms of latitude and longitude values. We prefer to use geolocation by choice, since, it permits farmers to precisely explore explicit areas in the field to monitor crop conditions, and participants get to know the definite location of banana growth. The banana sown date is also included, to compare with the harvest date and get to know whether the banana is grown in its full-time period, or not.

Next, the average temperature, humidity, and soil moisture data collected from the microprocessor are added. Along with the environmental data, the percentage of manures and pesticides used are added. This adds on to the confidence of consumers, to get to know how much chemically the banana is treated with during its growth.

At last, the harvest date, the type of storage method used to store unripened banana, the number of days it is stored before marketing, and the date it is ready to be sold. These details sum up to the harvesting and post-harvesting stages of banana, which is essential to know the quality of the banana. A smart contract: ishealthy is written based on the analysis of the banana plant growth details stored on the blockchain. Ishealthy is displayed as a Boolean value: true or false. The true value determines that the banana was grown under healthy environment, and hence the quality of banana fruit is good. The false value determines that the quality of banana fruit is not good.

Application Layer: A simple Web application is created to display the details of all the data stored in the blockchain. The data is fetched by using the farm license number as the key. The participant can provide the farm license number and pull in the details of banana grown in that farmland (Figs. 2 and 3).

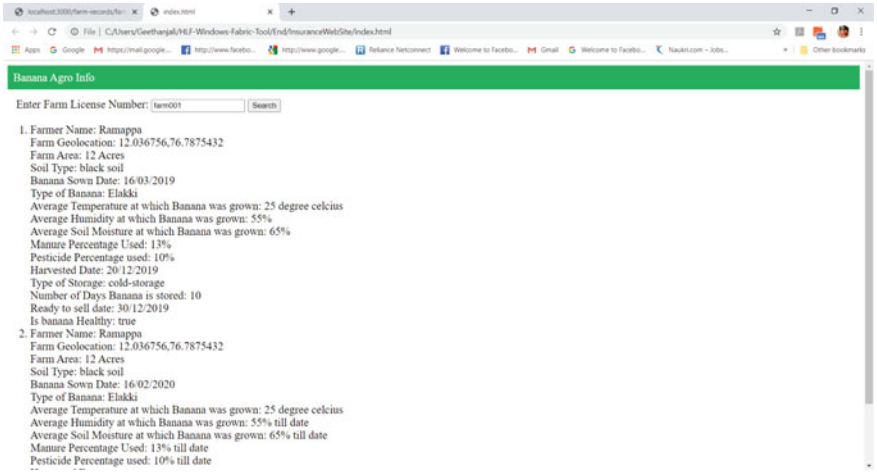


Fig. 2 Application layer of the Banana Agro System. The key farmer license number = farm001 is used here to fetch the details of banana growth

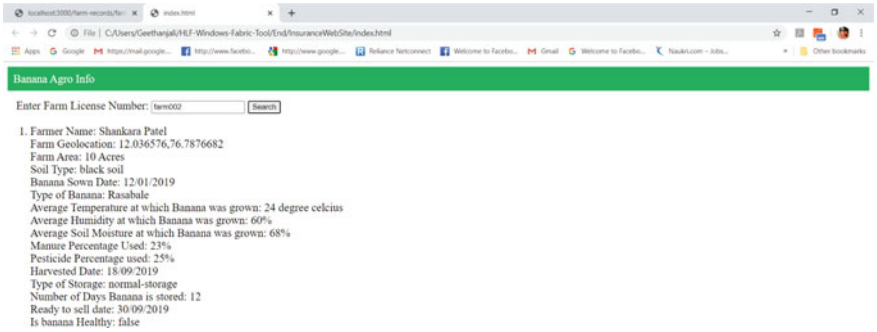


Fig. 3 Application layer of the Banana Agro System. The key farmer license number = farm002 is used here to fetch the details of banana growth

3 Benefits of the Blockchain Employing into Banana Agro System

- In the conventional method of data storage, data can be falsified. But blockchain will make it impossible to change any record, thus no participants can tamper the banana agro data stored on the blockchain. This helps farmers to deliver quality bananas to the right market for a decent price.
- Small farms will be able to attract more investments for their business.
- A permanent history of the banana agro data is created. This data can be used to increase yields of future banana plantations.

4 Conclusion

The pursuit of agricultural products is mainly to observe the environment of the product and to deal with problems with food safety. In this paper, a Banana Agro Model is proposed, which demonstrates how the quality of banana is assured using the blockchain technology. It provides accurate information on the path of banana growth, which is immutable and transparent. Participants can register and consult the blockchain ledger to be informed of the quality of the crop. The proposed system will help farmers decrease production costs and waste by optimizing the use of the inputs. Besides, it can increase yields by improving decision-making ability.

References

1. Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf
2. Kumar, M., & Iyenger, N.C. S. N. (2017). A framework for blockchain technology in rice supply chain management plantation. 125–130. https://doi.org/10.14257/astl.2017.146.22.
3. Xie, C., Sun, Y., & Luo, H. (2017). Secured data storage scheme based on block chain for agricultural products tracking. In *3rd International Conference on Big Data Computing and Communications (BIGCOM)* (pp. 45–50), Chengdu.
4. Bermeo-Almeida, O., Cardenas-Rodriguez, M., Samaniego-Cobo, T., Ferruzola-Gómez, E., Cabezas-Cabezas, R., & Bazán-Vera, W. (2018). Blockchain in agriculture: A systematic literature review. In R. Valencia-García, G. Alcaraz-Mármol, J. Del Cioppo-Morstadt, N. Vera-Lucio, M. Bucaram-Leverone (Eds.), *Technologies and Innovation. CITI 2018. Communications in Computer and Information Science, Vol. 883.* Springer: Cham. https://doi.org/10.1007/978-3-030-00940-3_4
5. Tian, F. (2016). An agri-food supply chain traceability system for China based on RFID & Blockchain technology. In *2016 13th International Conference on Service Systems and Service Management (ICSSSM)* (pp. 1–6). IEEE.
6. www.nhb.gov.in/report_files/banana/Banana.html.

Relative Study Between Technology to Perceive Hand Gestures

Chetan H., S. Praveen, S. Shreyas, Samridhi Singh, and R. Urvi

Abstract Interaction with each other helps human beings to share their intention, belief, understanding, and experience to individuals around them. This is not true in the case of people with hearing or speech disabilities. Communication for a hearing- or speech-impaired person is achieved through sign language. The main idea behind this problem is to develop a system for recognizing the signs, which would bridge out the communication gap between people with speech-impaired and normal people. The main goal of the paper is to track the region of interest (ROI) with the help of a camera and detect the target object by using object detection algorithms. Here, the target object is the hand and any gestures made by the hand are analyzed using different technologies.

Keywords Hand detection · Hand-gesture recognition · Sixth sense · Haar cascade · Tensorflow · CNN

1 Introduction

Hand motion acknowledgment framework got an extraordinary consideration over the last few years owing to its manifold applications and its capability to interact efficiently with machines via human and computer interaction. Wearable technology

H. Chetan (✉) · S. Praveen · S. Shreyas · S. Singh · R. Urvi
CMR Institute of Technology, Bengaluru, Karnataka, India
e-mail: chetan.h@cmrit.ac.in

S. Praveen
e-mail: spraveenpravi36@gmail.com

S. Shreyas
e-mail: shbglr06@gmail.com

S. Singh
e-mail: singhsamridhi21@gmail.com

R. Urvi
e-mail: rurvi09@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_62

is applied in various applications. Some people among us cannot communicate by talking. For enabling communication with them verbally, the sign language can be converted into voice directly.

There are different methods to develop a system that can be used for gesture detection. The major motivation behind our paper is to compare these technologies and analyze their working. The fundamental point of modeling hand signal acknowledgment framework is to make feature cooperation amidst human beings and computers where the apprehended motions can be used for managing or passing on essential data. The greater part of the analysts grouped motion acknowledgment framework into three stages subsequent to gaining the info image from the camera(s), recordings, or even information glove instrumented gadget [1]. These means are extraction method, highlights estimation and extraction, and order or acknowledgment.

2 Technology Used

2.1 Sixth-Sense Technology

Six sense is a wearable gestural connection that upgrades this present world around us with digital information and lets us use hand signs to interface with that data. It depends on the ideas of augmented reality and has well implemented its view. Sixth sense has incorporated these present reality objects with the advanced world. The remarkable sixth-sense innovation is a mixture of numerous advancements. It links technologies like hand motion acknowledgment, picture catching, handling, and control, and so forth. It superimposes the digital world in this present world. During the process of feature analysis and recognition, the greater part of the gesture recognition algorithm needs to remove features physically to complete the gesture recognition tasks, for example, support vector machines, undetectable Markov model, dynamic time regulation, and so forth. Random forest and artificial neural network.

Tensorflow is an open-source deep learning structure created by Google. Its front end supports numerous development languages, for example, Python, C++, and Java, etc. The backend is composed of C++, CUDA, etc. The algorithm executed in this structure can be handily transplanted on numerous heterogeneous frameworks [2]. It is supported by numerous designers. It can execute line and string tasks from the base layer, rapidly call equipment assets, give input information, graph node structure, and object function, and afterward, it can allocate nodes to numerous gadgets for equal activity. In this paper, the gesture recognition model is dependent on convolution neural systems which are built under the Tensorflow structure, and the essential network is designed. After the network parameters are optimized, the acknowledgment impact of the system is analyzed [3].

3 Methodology

3.1 Image Pre-processing

Image processing is a strategy to perform certain procedures on an image, in order to get an improved picture or to remove some supportive information from it. It is a kind of sign dealing with where data is an image and yield may be an image or characteristics/features identified with that image.

3.2 Segmentation

The principal objective of the division stage is to empty the foundation and commotions, leaving only the region of interest (ROI), which is the primary accommodating information in the image. This is accomplished by means of skin masking characterizing the edge on RGB construction and afterward changing over RGB shading space to dim scale picture. At last canny edge, the procedure is used to distinguish and recognize the presence of sharp discontinuities in an image, accordingly identifying the edges of the figure in the center [4].

3.3 Canny Edge Detection

Canny edge detection is a multi-step calculation that can perceive edges with commotion suppressed simultaneously. Smooth the images with a Gaussian channel to decrease noise and undesirable surfaces. It is chosen to the point that all edge components are kept while the greater part of the noise is suppressed (Fig. 1).

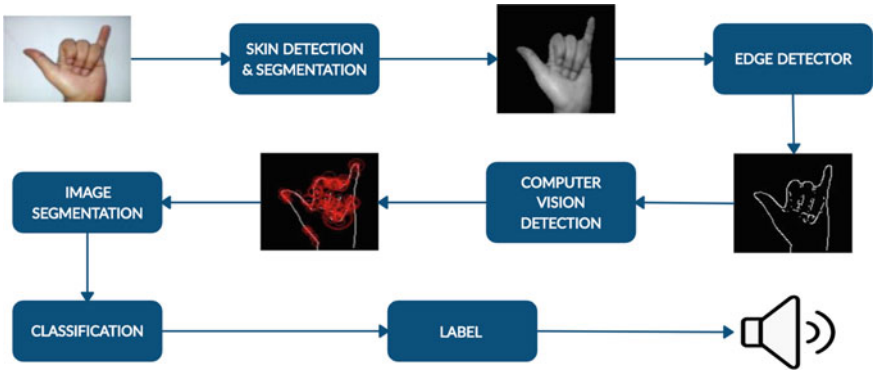


Fig. 1 Flow graph

3.4 SURF Feature Detection

Object recognition utilizing speeded-up robust features (SURF) comprises:

(i) Feature extraction—It is a significant advance in any item acknowledgment calculation. It alludes to the route toward removing significant data. The separated highlights must be an agent in nature, conveying significant and interesting traits of the picture.
(ii) Feature descriptors—It encodes interesting information into a series of numbers and goes about as a sort of numerical "fingerprint" that can be utilized to separate one component from another.
(iii) Feature matching—This method finds comparing features from two comparative images based on a search distance [4].

3.5 Sixth Sense

Components Used:

The tools are coupled in a pendant like an adaptable wearable gadget.
The equipment is a camera, projector, mirror, mobile component, and markers.

(A) Camera: Camera catches an object in view and keeps a tab on the client's hand motions.
(B) Projector: The projector presents perceptible data enabling surfaces around us to be utilized as interfaces.
(C) Mirror: Mirror reflects pictures onto the ideal surface.
(D) Smartphone: manages the sixth-sense programming, and controls the interrelation.
(E) Color marker: the marker is at the tip of the user's fingers. The improvements and approaches of these makers are deciphered into movements.

The technology used is:

(i) Gesture recognition
(ii) Computer vision
(iii) Radio Frequency Identification
(iv) Augmented reality

The disadvantage with the sixth sense is that, it has hardware limitations, continuous Internet connectivity cannot be provided to it, and also increased equipment cost.

3.6 HAAR Cascade

It is an AI-based strategy where coursework is set up from a ton of positive and negative pictures [5]. It is then used to perceive questions in various pictures.

The calculation has three stages:

(A) Haar feature selection
(B) Creating integral image
(C) Adaboost training
(A) Haar Features Selection

A Haar-like feature thinks about the area in a discovery window, summarizes the pixel intensity in each region, and figures the contrast between these sums.

(B) Creating integral image

Integral image is an image we get by the aggregate expansion of intensities on subsequent pixels in both horizontal and vertical axes. In image processing, we generally depend on features specific to a certain region of the whole image. Subsequently, we need properties of those particular regions.

(C) Adaboost training

Adaboost which chooses the best highlights and prepares the classifiers that utilize them. This calculation develops a "solid" classifier as a linear combination of weighted simple "weak" classifiers. One of the major issues with the haar cascade is processing time dependent on various classifier techniques because of this complexity will increase. Results are tricky to interpret. A superior result (contrasted with an ideal single classifier) will not generally be accomplished [6].

3.7 Convolutional Neural Networks (CNN)

A convolutional neural network is a deep learning computation that is utilized for picture classification, object acknowledgment, and detection. It can take in image data, distribute centrality (learnable loads and inclinations) to alternate points of view/inquiries in the image, and have the option to isolate one from the other. The prepreparing required in a ConvNet is a lot lower when contrasted with other gathering computations. While in unrefined systems, channels are hand—structured, with enough planning, and ConvNets can pick up capability with these channels/characteristics [7].

4 Results

A gesture is detected when the number of peaks in the histogram and the distance between them matches with an entry in a reference gestures library. Peaks in the histogram are detected as fingers. Gesture library matches the histogram using continuous Internet connectivity. This would be a disadvantage when it has to work in a remote location where Internet connectivity is not available (Fig. 2).

The following results were obtained by using Haar cascade algorithm (Fig. 3).

The hand gestures depicting a number of fingers are detected using OpenCV. Once the gesture is detected, the prediction is done using a trained Haar cascade model which is in.XML file extension.

The predicted output of the same is being displayed as a text on the screen. Audio output was also achieved for the same by using pyttsx library.

Fig. 2 Gesture recognition using sixth-sense technology

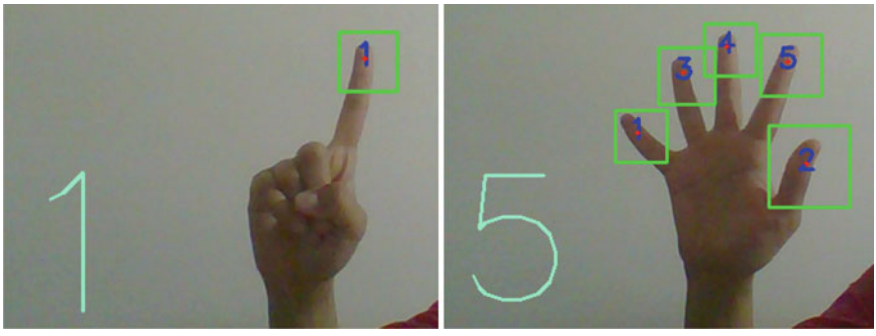


Fig. 3 Number recognition using cascade algorithm

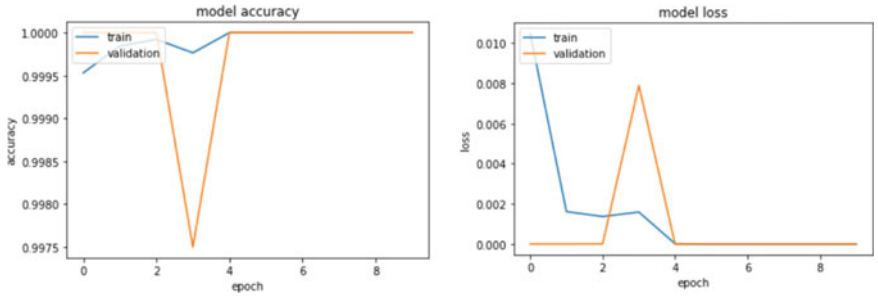


Fig. 4 Accuracy and loss of a training model

```
2000/2000 [==============================] - 4s 2ms/step
Accuracy:0.9990000128746033
```

Fig. 5 Tensorflow output

The pyttsx is a cross-platform text-to-speech library which is not platform dependent. The significant preferred advantage of utilizing this library for text-to-speech change is that it works offline.

Tensorflow output in Fig. 4 shows the graphical Keras model being trained with ten different types of images in which each image has 16,000 data. Before the model is trained, the data is divided into the ratio of 8:2 in which 80% is for preparing the model and the rest 20% is for validating a model. Also before training a model, the data needs to be made sure it is fit.

Once the model is trained, we would typically graph the accuracy of our model on the validation set, and choose a suitable number of epochs to train further to avoid overfitting. We might also consider introducing dropouts and regularisation.

However, we can see we are getting perfect accuracy on the validation set after just one or two epochs (Fig. 5).

5 Conclusion

This paper introduced a flexible methodology for gesture recognition. At first, the sixth-sense approach was utilized as it has inbuilt features to perform the required task, but owing to equipment limitations, the requirement of ceaseless power supply, and cost of the item, due to the following drawbacks, we then proceeded with the Haar cascade algorithm.

Utilizing the Haar cascade algorithm, the content and sound output were achieved. The fundamental issue that looked here was preparing the model. As training does

not occur constantly in the Haar cascade algorithm, training the model is not a simple task.

The final method was implemented using Tensorflow. The image was captured using OpenCV, and Raspberry Pi sends the image to Amazon Web services (AWS). The data fed to AWS sage maker is then handled by the Tensorflow library which detects and perceives gestures. CNN was utilized for picture classification, object recognition, and detection. Based on examination with the prepared model, the output is given through Raspberry Pi. This method has great accuracy compared to the above methods.

References

1. Deshmukh, S., Bhandaarkar, A., Chandak, S., Patil, S., Gaikwad, S., & Bhise, S. (2017). Hand gesture recognition system for blinds. Nagpur, India: Department of Computer Technology, Rajiv Gandhi College of Engineering & Research.
2. Zeng, Z., Gong, Q., & Zhang, J. (2019). CNN model "Design of gesture recognition based on tensorflow framework." In *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC).* https://doi.org/10.1109/itnec.2019.8729185.
3. Hussian, S., & Saxena, R. (2017). Hand gesture recognition using deep learning. In *2017 International SoC Design Conference (ISOCC).* https://doi.org/10.1109/isocc.2017.8368821.
4. Xu, Y., & Dai, Y. (2017). *Review of hand gesture recognition study and application.* South Korea: Department of Game and Multimedia Engineering PaiChai University. South Korea: Department of General education and International Trade Namseoul University.
5. Phuc, L. T. H., Jeon, H., Truong, N. T. N., & Hak, J. J. (2019, September). *Applying the Haar-cascade Algorithm for detecting safety equipment in safety management systems for multiple working environments.* Gyeongsang-si, Korea: Department of Chemical Engineering, Yeungnam University.
6. Alam, S. S., Islam, N. N. A. J., Ahammad, K. T. (2017, March) Hand gesture detection using haar classifier with appropriate skin color, kernel sizing & auto thresholding. *International Journal of Scientific & Engineering Research, 8*(3), 98. ISSN 2229-5518.
7. Sudharshan, D. P., & Raj, S. (2018). Object recognition in images using convolutional neural networks. In *2018 2nd International Conference on Inventive Systems and Control (ICISC).* https://doi.org/10.1109/icisc.2018.8398912.

Implementation of Electronic Records Management Systems: Potential and Challenges a Case Study of the Water and Power Development Authority (WAPDA) in Pakistan

Sidra Azmat Butt, Ingrid Pappel, and Karin Oolu

Abstract The state of record-keeping in the majority of local government organizations in Pakistan remains outdated where conventional systems are still followed for daily processes due to the vast amount of legacy systems and red-tapism. This research conducts a case study to understand the current state of record-keeping in a local government organization and the challenges in implementing an electronic records management system (ERMS) to promote the concept of good governance with the aid of transparency and accountability. A government department in Pakistan, i.e., the Water and Power Development Authority (WAPDA), is the selected case study; qualitative data were collected from expert interviews and employee questionnaires, which have been analyzed with tool-based thematic analysis. The proposed state of processes is displayed through a To-Be model by a visual representation of the advantages that an ERMS system can provide for a proper flow of steps and timely delivery of services.

Keywords Electronic record management system · ERMS · Information and communications technology · ICT · Information technology · Good governance

1 Introduction

In the digital world of today, the adaption of an electronic records management system (ERMS) is no longer a matter of choice, but a necessity. This is because of the radical penetration of computers into offices of public and state departments. IT systems

S. A. Butt (✉) · I. Pappel · K. Oolu
Information Systems Group, Tallinn University of Technology,
Akademia tee 15a, Tallinn 12618, Estonia
e-mail: sidra.butt@taltech.ee

I. Pappel
e-mail: ingrid.pappel@taltech.ee

K. Oolu
e-mail: karin.oolu@taltech.ee

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_63

based on computers have revolutionized the way of carrying out day-to-day tasks. This has resulted in the proliferation of data files in the system and, in turn, made it difficult to appraise and preserve these files [1]. Records maintained in a government organization are often the most crucial source for the processes of accountability. In a government organization such as the Water and Power Development Authority (WAPDA) of a developing country such as Pakistan, ERMSs have not been integrated nor been considered as part of the strategic management program. Therefore, this research analyses the current state of record-keeping in WAPDA, all the possible challenges that may be faced in its process of ERMS adaption and the potential for increasing efficiency in the daily workflow. A case study approach has been utilized along with interviews from officials and questionnaires from employees. This paper is part of the master thesis which was submitted and defended in 2019 at Tallinn University of Technology [2]. The outcome of this research can serve as practical knowledge for WAPDA with recommendations and redesigned "new connection request" as input for further improvement of the other processes in the organization.

The rest of the paper is organized as follows: in Sect. 2 related work as well as case is briefly described. Section 3provides the details of the methodology used for this research. Section 4 shows the results from the interviews and questionnaires, which are discussed along with recommendations in Sect. 5. And lastly, Sect. 6 presents the conclusion.

2 Related Work

2.1 Good Governance and ERMS

The idea behind the research paper is to facilitate the ideology of "Good Governance" in developing countries, through the application of necessary digital systems in government organizations. A modern form of government is about the efficiency of the system and accountability between the citizens and the government. According to Grindle, "the principle of accountability in the real sense, emphasizes answerability for the use of state resources and assets earmarked for specific purposes, subject to the laws and their requirements" [3]. An ERMS helps to build interoperable systems that lead to accountability [4]. Moreover, conformity to rules and regulations can be ensured through transparency as Nanda states, "in fact, the processes of decision-making, the ultimate decision reached and the government actions taken are expected to be made open and subject to check by other organs of government and non-governmental organizations" [5]. In the decision-making process, transparency is the free flow of information [6] that is easily accessible and understandable to the audience that may be affected by it. According to Nanda, "effectiveness is primarily the process of doing the right things, while efficiency is primarily the process of doing things right" [5]. A proper action plan should be developed by the government that takes care of the necessities of the societies while making sure that the resources are put for use in an efficient manner.

2.2 ERMS in Pakistan

Record management is defined as "the planning, controlling, directing, organizing, training, promoting, and other managerial activities involved in records creation, maintenance and use, and disposition in order to achieve adequate and proper documentation of the policies and transactions …and effective and economical management of its operations." [7]. The basis of any public sector organization is records management; this is because records are meant to be the "institutional memory" [8]. It is through effective implementation of an ERMS that the efficiency of a government in providing services to its citizens is increased [9].

The introduction of an ERMS in any organization brings a radical change in the organizational culture, the responsibilities, polices, and the standard procedures. The capability that an organization depicts in managing this change is the determining factor of the process of implementation [10]. According to the study conducted by Zinner and Andersen, the system that was followed throughout the Punjab Government departments was the one developed almost 80 years ago by the British [11]. With the growth in the government operations and the rapid increase in the responsibilities assigned to each department, paper records caused deterioration of the efficiency level as they were stored in big filing rooms with no system to track them and required immense manual work in case if a document had to be retrieved. There was no formal nor proper procedure for the record-keeping in the rooms, in the desks, or the offices where the records were stored [11]. The study further explains that the major constraint was the lack of technical abilities in the IT Department of Punjab. There is near to none experience for such initiatives in the country. The staff is trained to follow manual procedures and shifting to the electronic system means providing them with the technical skills and training them accordingly [11].

The implementation of ERMS has remained limited in Pakistan even though it has increased the level of efficiency and effectiveness in a few government departments such as NADRA [12]. According to Yahaya, the workstations in most of the government departments in Pakistan generate a huge volume of documents and emails [13]. The issue here is to be able to manage these volumes of electronic documents and records. Moreover, to implement an ERMS, some important factors need to be analyzed so that the records are obtained, created, and also managed as per the policies of the department. The factors that may pose as challenges to the ERMS implementation are presented in Table 1.

2.3 Case Background

WAPDA was created in the 1958 and has the responsibility to coordinate and give out unified directions for the development of plans related to the Power and the Water Sector. In 2007, the organization was split where Pakistan Electric Power Company (PEPCO) was made responsible for the production of power, its supply,

Table 1 Challenges in ERMS implementations

IT literacy	No formal IT curriculum in educational institutes. Lack of training and skillset further translated into the workforce of an organization
E-readiness	Tools that aid in assessing the e-readiness that can be divided into two main categories: one that assesses the basic infrastructure of the country and the second that assesses the ability that a nation possesses in order to take advantage of the ICTs
IT infrastructure	An IT infrastructure means the technical bases like hardware, the communication technology, software, basic applications, and the capability of the user to use these applications [14]
Co-operative Management	If there is no support from the management, it may end in preventing the users to use the particular system [15]. With the implementation of IT, the power enjoyed by the top management is challenged, which is why many times they are reluctant to digitalize the system [16]
Training	Training can be done by introducing in-service programs, workshops, conferences, etc., that cover the basic principles and functionalities of an ERMS [17]. These trainings can be time-consuming and expensive for the workforce that is not technologically advanced
IT Policy	It is imperative to have policies in place for organizations to take digitalization seriously [18]. Different countries have different policies to ensure the safety of data as well as to prove the authenticity and reliability of the records [19]
Culture	Hofstede define culture as "patterned ways of thinking …the essential core of culture consists of traditional (i.e., historically derived and selected) ideas and especially their attached values" [20]. In Pakistan, paper is considered as a trusted form of documentation against electronic records

transmission, and billing, whereas WAPDA was given the responsibility of water and hydropower development. The main objective of the organization is to make sure that the supply of electricity is reliable throughout the country. As per the statistics of 2012, WAPDA has 20.12 million users with a major share of 43% in the domestic sector, 26% in the industrial sector, and approximately 12% in the agriculture sector [21]. WAPDA has monopolized the distribution and the transmission networks. When WAPDA was unbundled, there were nine distribution companies (DISCOs) created, four of the generation companies (GENCOs), and a transmission company. After the transmission, power is then handed over to the eight distribution companies (shown in Fig. 1) in order to make sure that it is evenly distributed amongst the consumers of the state. For the purpose of the study, the services provided by Gujranwala Electric Supply Company (GEPCO) will be discussed for a thorough understanding of the operations in the department, translating to WAPDA eventually.

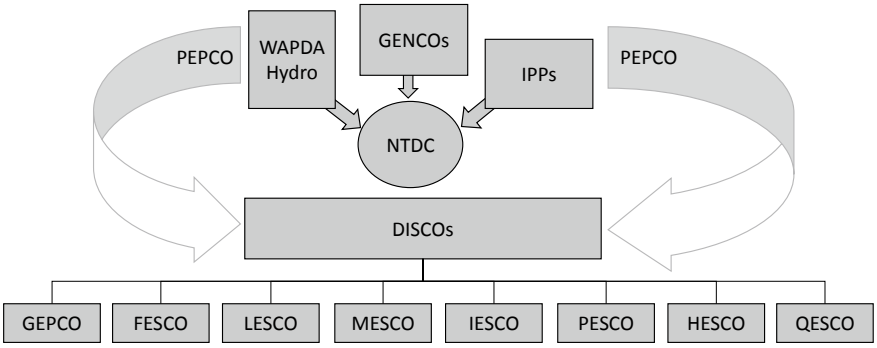


Fig. 1 Divisions and subdivisions of WAPDA

3 Research Methodology

This research utilizes the case study approach in order to understand the current state of record-keeping in the country and the challenges in implementing an ERMS. The case study for this paper is WAPDA, a government department in Pakistan. Data in this case were collected from employees at different levels of the department in order to determine the current situation of record-keeping and services provided. Interviews were conducted from the following officials at GEPCO:

- Sub-divisional Officer (SDO)
- Executive Engineer (ExEn)
- Customer Service Specialist (CSS).

Interviews (approx 40 min long) were then analyzed using an open-source qualitative data analysis software, RQDA to carry out a thematic analysis. On the other hand, a questionnaire was also used in order to gather information from the employees of WAPDA and its subdivision GEPCO. A total of 50 employee responses were recorded.

4 Results

4.1 Interview Results

Different codes that were generated during the analysis using RQDA are shown in Fig. 2. These codes were then used to generate themes which are shown in Table 2.

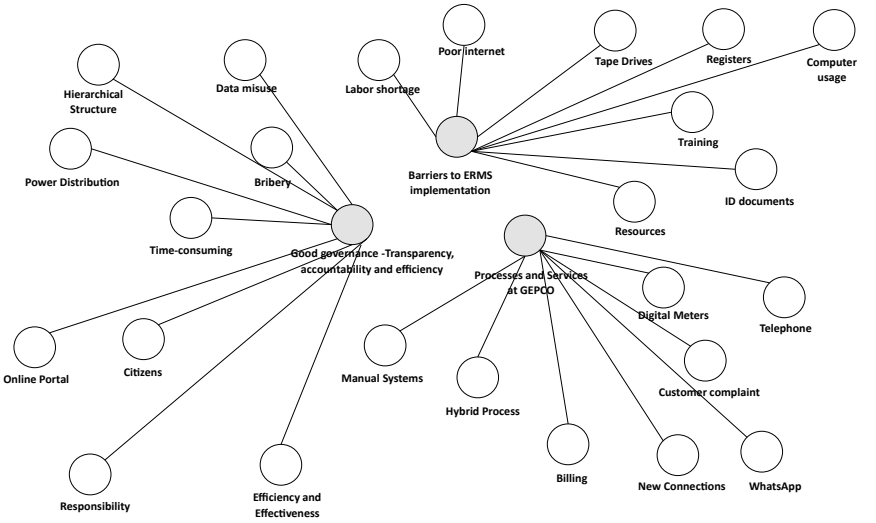


Fig. 2 Thematic map

4.2 Questionnaire Results

All of the employees agreed to the idea that technology in this era has become an essential part of our daily routine. The responses to the main issues that currently WAPDA faces in its daily workflow were mixed where most of the respondents agreed that the lack of material resources (70%) and skilled labor (62%) was to blame. 72% of the respondents claimed that poor customer support was the reason for the WAPDA's reputation to deteriorate. The majority of employees believed that the method of record-keeping can be improved and that ERMS can serve as an effective solution in order to increase the efficiency. While most of the respondents claimed that through the implementation of ERMS there will be transparency of processes and is essential to increase efficiency, many believed that it is an expensive process and that unskilled labor might lose their jobs in the organization. On the required IT infrastructure, 52% respondents believed that the necessary infrastructure is not present. 54% of the respondents agreed that WAPDA employees do not have any sort of IT skills which may become a hindrance to the ERMS implementation and therefore, 68% agreed that there will be resistance to this change.

5 Discussion and Recommendations

The research suggested that to effectively enroll an ERMS in a government department such as WAPDA, and it is of paramount importance to first analyze the existing hybrid system for records management. For an effective ERMS implementation,

Table 2 Themes

Processes and Services at GEPCO	• New connections request: – Office visit – Website Even after a new connection request is filled online, the rest of the process remains manual where most of the communication between the team is through WhatsApp messages or telephone calls • Complaint and Customer Services: – Online portal – Manual entry in register These complaints travel from the clerical staff to the SDO and then to the ExEn, where documents are exchanged through postal service. If the complaint is not redressed, customer has to pay multiple visits to the office and every time his complain is fetched from a register that can be time-consuming • Billing Services: Lineman or staff take meter readings through their phones, which are then communicated through WhatsApp messages and then updated into the systems manually in order to generate manual as well as online bills
Barriers to ERMS implementation	• Shortage of labor which also shows that the processes are redundant which means a large workforce is required (SDO) • Difficult to process the new connection requests, or cater to the complaints of the individuals when the necessary equipment is not readily provided by the state (SDO) • Most of the workers have been working the same job since more than three decades where no training is provided to them in order to improve their skill set as per the current requirements (ExEn) • Power is distributed through hierarchy that often leads to misuse of power and authority. Bribery, corruption, and nepotism (SDO) • Citizens of the state unaware of their civic responsibility where even the online portals are exploited in order to register false complaints and create issues for a department which already is short of workforce and resources (CSS) • Internet was only available to the managerial level of the department that was used seldom to send emails. ID documents shared through postal service (ExEn) • No archival repository besides a tape drive and registers (SDO) • The clerical staff is barely equipped to use a computer, and most of the details and complaints are still jotted down in a register (CSS)
Good governance: Transparency, accountability, and efficiency	• Follow up on a complaint requires periodic visits to the office or phone calls. Staff already overworked, do not cater to each inquiry, and are often found rude while providing inadequate responses. This has slowed down the processes, created undue frustration, encourages bribery culture and resentment of the common public towards the government departments in general and WAPDA or GEPCO in particular (CSS) • Hybrid system (paper + online) is being used for record-keeping. GEPCO uses tape drives in order to save all the data including the sensitive information. There is no cloud-based system or a database that saves the data collected from the public. Tape drives are not reliable, and there can be a huge loss of data due to any natural calamity, technical faults, or personal favors (SDO) • Document search is time-consuming, “once we took 5 h to search a document. It was some bureaucrat’s complain so we had to” (CSS)

GEPCO needs to update its existing computer infrastructure, introducing systems that are up-to-date and networks that will be able to support the ERMS. Different servers such as software, processing power have to be determined that could be used for an ERMS implementation; according to the department requirements and server storage, capacity has to be defined. According to the interviews, Internet connection was poor and barely used to exchange emails. The documents exchanged over emails also have to be analyzed to distinguish which emails are considered as records and then a records management policy has to be devised to retain email records. The highest bandwidth of network servers as well as multifunctional devices has to be put in place to accommodate high volumes of records.

Furthermore, the management needs to be trained beforehand, to make sure that they are capable enough to handle the change that an ERMS system will bring to the organization. If workers are included in the ERMS implementation process and planning, it is more likely that they are willing to get training as well as accept change. Moreover, with the ERMS implementation, not only the work procedures will change but also the organizational culture. If an ERMS is used, it can eventually lead to accountability and transparency which means power will be redistributed and may shift down to the system handlers and users. It will have an immense effect on the decision-making process when the system takes control to assess the workload and workflow [11]. Pakistan has a lot of potential and many trained professionals in the IT department who understand the need of implementing the ERMS in various government departments. The need here is to make sure that they have enough resources and finances in order to pursue their ideology and introduce the system of digitization in at least the major departments run by the government such as WAPDA. Pakistan can take inspiration from Estonia as "Best Practice" title was won for its transition to a paperless management system in 2009 [22]. The Government of Estonia has always stayed a step ahead and provided the system with a "defined regulations on standard, and pros, procedures, methods, and products to assist the modern IT-evolution processes" [22]. Moreover, an X-road has been established and interoperability has been created to strengthen private–public relationship [23] and eID is being used which further compliments digital signatures [24], e-invoicing [25], e-voting, etc. Therefore, there are several factors that need consideration in order to make sure that the implementation of ERMS in a government organization is successful which includes:

- Increasing the IT literacy within the organization.
- Determining the e-readiness.
- Developing a reliable IT infrastructure.
- Having a co-operative management.
- Developing training programs for employees to equip them with required IT skills.
- Formalizing IT policies to ensure that the implemented ERMS is utilized as intended.

Each one of these factors has to be systematically researched further to understand the needs and expectations of different stakeholders involved in the ERMS implementation in WAPDA and other government organizations in general.

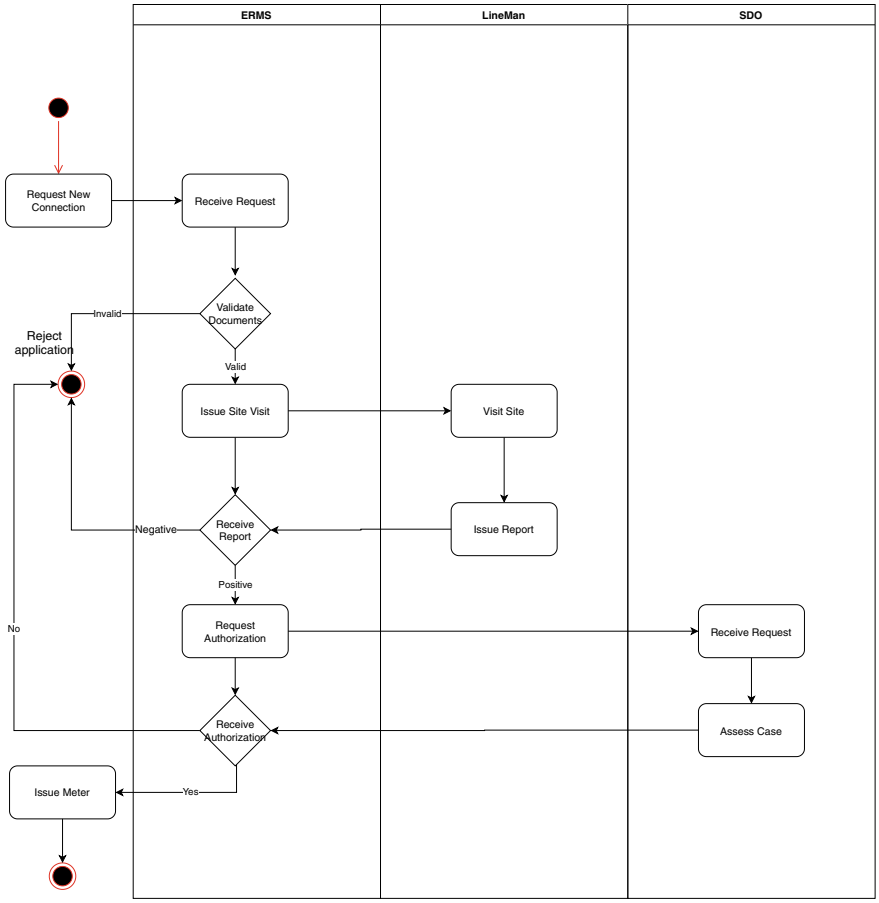


Fig. 3 To-be model for new connection request at GEPCO

The purpose here is to design a service that is based on the experiences from the customer-end perspective. By engaging all of the stakeholders in the process of service design, it results in a system that is cohesive and aligns well with the needs of the customers providing them with a better service. Therefore, the To-Be model, depicted in Fig. 3, has been designed in order to help visualize that the implementation of an ERMS that can remove the reasons behind delayed services and the iterations that are redundant. The process for new connection request at GEPCO has been re-designed to shorten the workflow and subsequently to increase efficiency.

6 Conclusion

The purpose of the paper was to analyze the enrolment of an ERMS in a government department in order to ensure efficient processes and services which facilitate good governance, while also analyzing the possible challenges that might arise. It was found out that the implementation of an ERMS and a better management of public information is capable of increasing transparency and accountability. The lack of IT infrastructure and limited resources is identified as the root cause of outdated record management in WAPDA; therefore, the need for reforms and implementation of an ERMS cannot be overemphasized. The ERMS implementation can be tedious, expensive, time-consuming, and cause resistance to change; but it is important to realize that in order to survive in this technologically advanced world, an ERMS is an essential first step towards digitalization.

References

1. Pappel, I., Butt, S. A., & Draheim, D. (2020). On the specific role of electronic document and records management systems in enterprise integration. In *Proceedings of ICICT 2020—the 5th International Congress on Information and Communication Technology*. Springer.
2. Butt, S.A., Pappel, I., & Oolu, K. (2019). *Electronic Records Management and its Effectiveness in the Local Government of Pakistan*. Master's thesis, Tallinn University of Technology, Tallinn.
3. Grindle, M. (2004). Good enough governance: Poverty reduction and reform in developing countries. *Governance*, *17*, 525–548.
4. Paide, K., Pappel, I., Vainsalu, H., & Draheim, D. (2018). On the systematic exploitation of the Estonian data exchange layer X-Road for strengthening public-private partnerships. In *Proceedings of ICEGOV'2018—the 11th International Conference on Theory and Practice of Electronic Governance* (pp. 34–41). ACM.
5. Nanda, V. (2006). The "good governance" concept revisited. *Annals of The American Academy of Political and Social Science*, *603*, 269–283.
6. Graham, J., Plumptre, T. W., & Amos, B. (2003). Principles for good governance in the 21st century. In *IOG Policy Briefs*. Institute on Governance (IOG).
7. Mutimba, C. J. (2014). *Implementation of Electronic Document and Records Management System in the Public Sector: A Case Study of the Ministry Higher Education Science and Technology*. Master's thesis, University of Nairobi, Nairobi.
8. Pollitt, C. (2000). Institutional amnesia: A paradox of the 'information age'? *Prometheus*, *18*, 5–16.
9. Pappel, I., Tsap, V., & Draheim, D. (2019). The e-LocGov model for introducing e-governance into local governments: An Estonian case study. *IEEE Transactions on Emerging Topics in Computing*, 1
10. Gunnlaugsdottir, J. (2009). The human side of ERMS: An Icelandic study. *Records Management Journal*, *19*, 54–72.
11. Henriksen, H., & Andersen, K. N. (2008). Electronic records management systems implementation in the Pakistani local government. *Records Management Journal*, *18*, 40–52.
12. Bokhari, H., & Khan, M. (2012). Digitisation of electoral rolls: Analysis of a multi-agency e-government project in Pakistan. In *Proceedings of ICEGOV'12—the 6th International Conference on Theory and Practice of Electronic Governance* (pp. 158–165). ACM.
13. Yahaya, D. H. Z. (2002). Keynote address: National archives of Malaysia. *ESARBICA Journal*, *21*, 63–68.

14. Husain, I. (2000). *Pakistan: The economy of an Elitist State*. Oxford: Oxford University Press.
15. Simon, J. C. (2015). Building your own ERMS. *Library Hi Tech News*, *32*, 9–12.
16. Marilyn, S. (2005). *Traces: Document, record, archive, archives* (pp. 1–20). Centre for Information Studies.
17. Ashcroft, L., & Watts, C. (2005). ICT skills for information professionals in developing countries: Perspectives from a study of the electronic information environment in Nigeria. *IFLA Journal*, *35*, 6–12.
18. Avgerou, C. (2010). Discourses on ICT and development. *Information Technologies & International Development*, *6*, 1–18.
19. Gunnlaugsdottir, J. (2006). *The Implementation and Use of ERMS: A Study in Icelandic Organizations*. Ph.D. thesis, University of Tampere.
20. Hofstede, G. (2001). Culture's consequences: Comparing values, behaviors, institutions and organizations across nations. *Behaviour Research and Therapy*, *41*
21. Malik, A. (2007). *Effectiveness of Regulatory Structure in the Power Sector of Pakistan*. PIDE-Working Papers, 25
22. Pappel, I., & Pappel, I. (2011). Methodology for measuring the digital capability of local governments. In *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance. ICEGOV '11* (pp. 357–358), New York, NY, USA. Association for Computing Machinery.
23. Paide, K., Pappel, I., Vainsalu, H., & Draheim, D. (2018). On the systematic exploitation of the Estonian data exchange layer X-road for strengthening public-private partnerships. In *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance. ICEGOV '18* (pp. 34–41), New York, NY, USA. Association for Computing Machinery.
24. Pappel, I., Pappel, I., Tepandi, J., & Draheim, D. (2017). Systematic digital signing in estonian e-government processes—influencing factors, technologies, change management. *Transactions on Large-Scale Data- and Knowledge-Centered Systems*, *36*, 31–51.
25. Pappel, I., Pappel, I., Tampere, T., & Draheim, D. (2017). Implementation of e-invoicing principles in Estonian local governments. In *Proceedings of the 17th European Conference on Digital Government (ECDG): Military Academy, Academic Conferences and Publishing International Limited* (pp. 127–136).

Machine Learning Based-Approach on Predicting Suicidal Ideation

Moname Majumder, Anika Tabassum, Jenifa Mousumi, and Mohammad Rezwanul Huq

Abstract Suicide is an immensely important and critical issue in today's world. Sometimes, it takes a suicide attempt for other people to realize that a person they know might have suicidal tendencies. Identification of suicidal ideation in time can prevent many untimely deaths. When it comes to expressing the sentiment, Facebook is one of the most viable mediums. But information from Facebook alone cannot paint a clear picture of one's psyche. Some standardized psychological tests for identifying depression, hopelessness, anxiety, and even suicide can help to understand one's mental state better. In this paper, we have used several informative sets of features, including demographic characteristics, Facebook status, and psychological test responses to predict suicidal ideation. To do so, we have created three different data models from our dataset and implemented several well-known, widely used machine learning algorithms on them. Out of all the different models, the Naïve Bayes algorithm implemented on a data model containing demographic data and psychological test responses produces the best result.

Keywords Suicidal ideation · Machine learning · Beck's depression inventory II · Beck's hopelessness scale · The suicide behaviors questionnaire-revised · Columbia suicide severity rating scale

M. Majumder (✉) · A. Tabassum · J. Mousumi · M. R. Huq
East West University, Aftabnagar, Dhaka 1212, Bangladesh
e-mail: monamemajumder@hotmail.com

A. Tabassum
e-mail: tabassum.anika24@gmail.com

J. Mousumi
e-mail: jenifa.mousumi@gmail.com

M. R. Huq
e-mail: mrhuq@ewubd.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_64

1 Introduction

An alarmingly large number of people across the world attempt suicide every year. If identified in time, people possessing suicidal tendencies may be helped. Many types of researchers are going on and trying to detect suicidal ideation and behavior. Many factors are responsible for suicide, which are called suicidal factors. Some of them are personal issues, such as hopelessness, severe anxiety, depression, chronic pain, and medical illness. The others are social factors, like social isolation or negative life events, including traumatic events, physical illness, affective disorders, and previous suicide attempts [1]. Numerous scales have been developed [2] to identify the suicidal intensity of human beings. With the growing and advanced internet technologies, the tendency of sharing personal feelings, day-to-day stories, thoughts have increased profoundly. These online contents could help detect an individuals' suicidal ideation. Some of the researchers have used posts from various popular social networking Web sites along with demographic data as other features and analyzed both. But those studies have their limitations. In this paper, social media (Facebook) posts along with psychological tests about depression, hopelessness, and suicidal ideation have been used to get a better picture of a participant's psychological state. With help of these attributes, even without the help of a psychiatrist or a clinical record, a participant's suicidal behavioral status can be predicted.

2 Related Work

The techniques of research used for examining suicide also covers numerous methods and fields. Example: A modest mental survey (MMS) is conducted by the Taiwan Suicide Prevention Center (TSPC), which has calculated suicidal inclination with 85 percent accuracy designing a system on Facebook [3]. The paper mainly analyzes suicide data, identifies demographic data and important traits that lead to suicide attempts, and predicts potential attempts with reasonable accuracy [4]. The research teams made use of 57 characteristics from a psychiatric hospital's self-harm surveillance report to compare ID3, C4.5, and Naïve Bayes and detect the attributes of a person with suicidal thoughts by using data mining techniques to commit suicide attempts [5], repetitively. The percentage of suicides for each one of the Indian states and the demographic characteristics of that state have been used as attributes, and the three kinds of the dataset are census data, marital status data, and state educational status data. The Pearson correlation was calculated to assess the impact of the attributes on the suicides number and linear regression model for suicides predicted number [6]. Utilizing machine learning to electronic medical records in a broad clinical database, the researchers have designed machine learning models that predicted accurately potential suicide attempts [7]. This study presents an approach to predict suicidal ideation through Facebook status, demographic data, and multiple-choice self-report inventory. Though there have been some attempts at predicting suicidal

behavior using psychological tests, the machine learning approach was not applied [8].

Standard psychological tests for identifying depression, hopelessness, and anxiety which were used in this paper are given below.

Beck Depression Inventory: There are three versions—BDI, BDI-1A, and the BDI-II. In this study, BDI-II has been used. BDI-II features 21 items [9] to assess signs and extent of depression in individuals.

Beck Hopelessness Scale: The Beck Hopelessness Scale [10] offers a measure of self-reporting of one's pessimistic assumptions about the future. It includes 20 true-false statements.

The Suicide Behaviors Questionnaire-Revised (SBQ-R): The Suicide Behaviors Questionnaire-Revised (SBQ-R) [11] is a self-report psychological questionnaire. On the SBQ-R, the highest score of eighteen is probable.

Columbia Suicide Severity Rating Scale: The Columbia Suicide Severity Rating Scale (C-SSRS) [12] is a method of assessment that assesses behavior and suicidal ideation. The C-SSRS is reliable and valid in the identification of suicide risk in several research studies. Research into using the Columbia protocol has verified the validity and reliability of the protocol for risk assessment [13]. C-SSRS has been used to create the **ground truth** of this research as its predictive validity, sensitivity, and specificity are very high [14]. The dataset was labeled based on this scale, so scoring was done very cautiously [15, 16].

3 Dataset Preparation

A survey was formed to gather assessments from several people through Google Forms. In the survey, respondents were told to take standardized multiple-choice self-report inventory, followed by many other demographic questions. Participants answered four psychological test questionnaires. They are—(1) Beck Depression Inventory, (2) Beck's Hopelessness Scale, (3) The Suicide Behaviors Questionnaire-Revised (SBQ-R), and (4) Columbia–Suicide Severity rating Scale. The respondents might also choose to share zip files or Facebook usernames for Facebook posts. There are 99 individuals participated in the survey between February 2020 and April 2020. Noisy responses were removed, resulting in a group of 85 users. The set was composed of 47 males and 38 females. In this way, two user classes were formed: a positive class of 20 users scored significantly higher for suicidal ideation, and a negative class of 64 users, those with little to no symptoms of the condition.

4 Proposed Methodology

The class label, i.e., ground truth, has been defined using the Columbia Suicide Severity Rating scale while the other scales have been used as features of our model. Based on the dataset, three data models have been created.

Data model 1: This data model consists of—(a) Facebook status (b) prediction class (The predicted class has been defined using Columbia Suicide Severity Rating Scale.).

Data model 2: This data model consists of—(a) responses of the demographic data (b) answers to the questions of Beck's depression inventory II, Beck hopelessness scale, The Suicide Behaviors Questionnaire-Revised (c) cut-off results of the above-mentioned three psychological scales (d) prediction class.

Data model 3: This data model consists of—(a) Facebook status (b) responses of the demographic data (c) answers to the questions of Beck's depression inventory II, Beck hopelessness scale, The Suicide Behaviors Questionnaire-Revised (d) cut-off score of these psychological scales (d) prediction class.

As preprocessing, the textual data was processed by performing data cleaning, stemming, and finally creating a sparse matrix using the Bag of words model. For features other than the Facebook status, categorical data were converted into numeric values as data preprocessing. Then, on three models, six machine algorithms were applied KNN, Logistic Regression, Support Vector Machine, Naïve Bayes, Decision Tree, and Random Forest. In Fig. 1, the different methods of this research have been visualized.

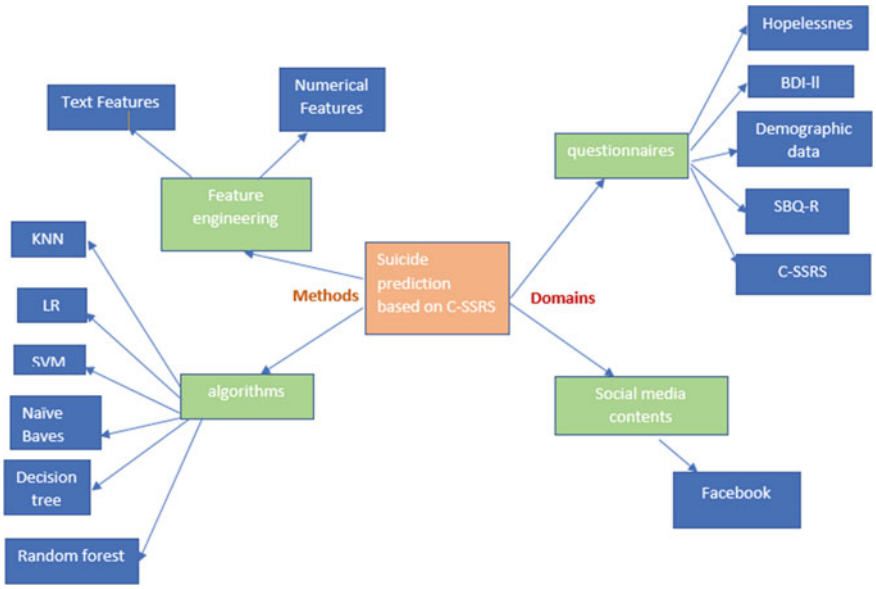


Fig. 1 Categorization of suicide prediction: methods and domains

5 Performance Evaluation

The Columbia suicide severity rating scale has been used to define the ground truth. Figure 2 shows, among all the participants, 24% had suicidal ideation, and 76% did not have it. This was computed based on the Columbia suicide severity rating scale.

In data model 3, we have applied tenfold cross-validation and calculated different performance measures(accuracy, precision, recall, and F1 score) for six machine learning classification algorithms.

In Fig. 3, we can see that in most of the cases, the values of the performance measures are near perfection. This high accuracy, precision, recall, and F1 score indicate that the model is extremely overfitted. Here, Gaussian Naïve Bayes algorithm seems to be underperforming for this data model. It is because it does not incorporate Laplacian Smoothing. But, when we implement multinomial Naïve Bayes, we see that there is a visible improvement in the performance. It is because multinomial Naïve Bayes considers Laplace Smoothing and prevents the model from setting null probabilities when the frequency is zero. Due to its visible overfitting nature of this model, this approach with all the features seems unsuitable.

We have applied the six classification algorithms on data model 1 as well. Figure 4 shows the performance of data model 1 that is remarkably weak compared to the performance of data model 3. Accuracies of none of the models exceed 70%.

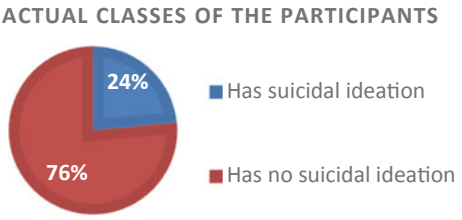


Fig. 2 Distribution of class based on the ground truth

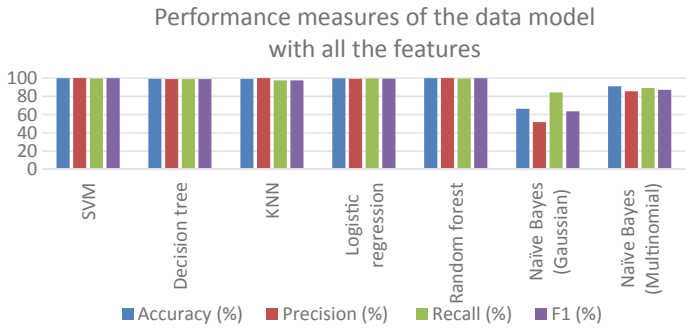


Fig. 3 Accuracy, precision, recall, F1 scores for data model 3

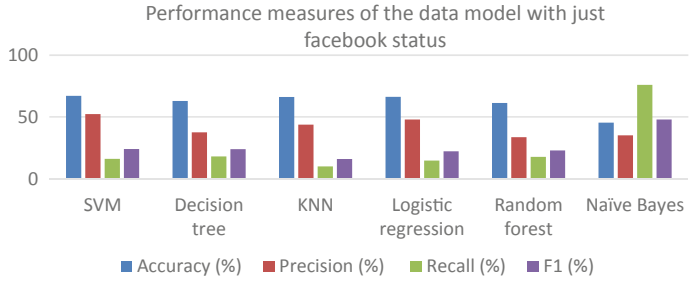


Fig. 4 Accuracy, precision, recall, F1 scores for data model 1

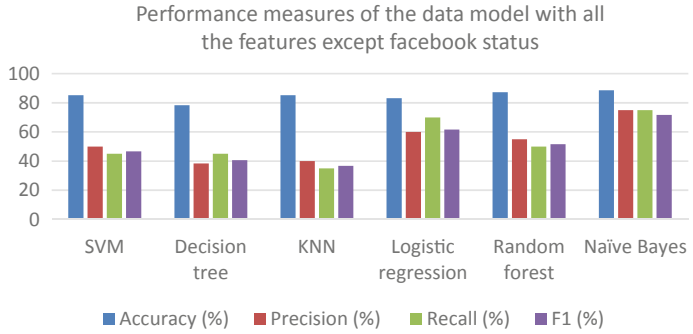


Fig. 5 Accuracy, precision, recall, F1 scores for data model 2

However, most of these models seem to perform worse when we analyze precision, recall, and F1 scores. It is because accuracy uses both TN (true negative) as well as TP (true positive) as numerator while the other measures use only TP. The values of recalls are significantly low, which must be an effect of high FN (false negative). From Fig. 4, it is clear that using just the text data to predict suicidal ideation results in underfitting. Interestingly, Naïve Bayes has low accuracy but high recall and precision when compared to the other model.

From Fig. 5, it is clear that when it comes to precision, recall, and F1 score, SVM, KNN, and decision tree display inferior performance. The performance of Random Forest is also flawed. The performance of logistic regression though is not bad but is not very good either. Out of the six algorithms, Naïve Bayes has the best performance for every performance measure.

6 Discussion

Figure 3 shows that the models are overfitted. These classification models are a combination of data model 1 and 2. Even though the model with just the Facebook status had a feeble performance, combining it with the model of other features resulted in a near-perfect performance. The reason behind this is that while combining models 1 and 2, there was a repetition of a number of features. Data model 3 was constructed by taking the Facebook status of a participant and adding it with his other features. So if someone had 'n' number of status, his other features would be repeated 'n-1' times. Excluding the column for Facebook status, there are still 55 columns of other features left, which are repeated 'n' times for each participant. Because of this redundancy, the model got overfitted and resulted in a near-perfect performance. Hence, the model with all the features (data model 3) is not suitable based on our understanding.

In our research, we are trying to predict an individual's suicidal ideation. For this study, the FN (false negative) cases carry great significance. If a person with a suicidal tendency is identified as a person with no suicidal ideation, it could be disastrous. So while considering the performance measures, examining FN is very important. Accuracy considers FN, but it is overshadowed by a large number of TN (true negative) cases since our dataset contains 76% of non-suicidal participants, and our models can identify the TN cases quite accurately. Precision does not consider FN cases at all. But recall takes FN cases into account. As shown in Fig. 4, the value of recall is significantly low; it can be concluded that that model contains a considerable amount of FN cases. Apart from the underfitting, this is another reason why data model 1 is not useful.

Neither precision nor recall considers both FN (false negative) and FP (false positive) cases. So, to have a balance, we have calculated the F1 scores, which include both FP and FN. For an unbalanced dataset like ours, the F1 score is a useful metric.

From the analysis of Figs. 3, 4, and 5, it is clear that the model with all the features is too overfitted, and the model with just status is also under fitted. Compared to these two extremes, data model 2 produces relatively good results. From Fig. 5, it has been analyzed that Naïve Bayes performs best for the data model 2. Hence, it can be concluded that for a new test data, the implementation of Naïve Bayes on data model 2 will produce the best result.

7 Conclusion and Future Work

In this paper, we have predicted the suicidal ideation of an individual based on their demographic data, Facebook status, and their responses to some well-established psychological scales. From different combinations of these data types, we have created three data models and implemented six machine learning algorithms on each

of them. After analyzing these different models, we can conclude that the Naïve Bayes classification algorithm implemented on the data model containing all the features except the Facebook status (data model 2) produces the best performance when it comes to predicting suicidal ideation.

From the performance of data model 1 and 3, it is evident that there is work to be done on the data models that deals with text data. In the future, we would like to extend this research by working on natural language processing. We can also continue this study to determine which psychological scale has the most impact on the outcome.

References

1. Kiriakidis, S. P. (2015). Elderly suicide: Risk factors and preventive strategies. *Annals of Geriatric Medicine and Research*, *2*(2), 1028–1032, [Online]. Available https://www.jscimedcentral.com/Gerontology/gerontology-2-1028.pdf.
2. Ghasemi, P., Shaghaghi, A., & Allahverdipour, H. (2015). Measurement scales of suicidal ideation and attitudes: A systematic review article. *Health Promotion Perspectives, 5*(3), 156–168. https://doi.org/10.15171/hpp.2015.019.
3. Chiang, W. C., Cheng, P. H., Su, M. J., Chen, H. S., Wu, S. W., & Lin, J. K. (2011). Socio-health with personal mental health records: Suicidal-tendency observation system on facebook for Taiwanese adolescents and young adults. In *2011 IEEE 13th International Conference e-Health Networking, Application & Services (Healthcom), 2011* (Vol. 50, pp. 46–51). https://doi.org/10.1109/health.2011.6026784.
4. Agarwal, T., Dhawan, A., Jain, A., Jain, A., & Gupta, S. (2019). Analysis and prediction of suicide attempts. In *2019 International Conference Computing Power Communication Technologies* (pp. 650–665).
5. Boonkwang, K., Kasemvilas, S., Kaewhao, S., & Youdkang, O. (2018). A comparison of data mining techniques for suicide attempt characteristics mapping and prediction. In *Proceedimgs of—2018 International Seminar on Application for Technology of Information and Communication Create Technolgy Humum Life, iSemantic 2018* (pp. 488–493). https://doi.org/10.1109/isemantic.2018.8549835.
6. Priyanka, S., Galgali, S., Priya, S., Shashank, B. R., & Srinivasa, K. G. (2017). Analysis of suicide victim data for the prediction of number of suicides in India. In *2016 International Conference on Circuits, Controls, Communications and Computing I4C 2016* (pp. 1–5). https://doi.org/10.1109/cimca.2016.8053293.
7. Walsh, C. G., Ribeiro, J. D., & Franklin, J. C. (2017). Predicting risk of suicide attempts over time through machine learning. *Clinical Psychological Science, 5*(3), 457–469. https://doi.org/10.1177/2167702617691560.
8. Lester, D. (1970). Attempts to predict suicidal risk using psychological tests. *Psychological Bulletin, 74*(1), 1–17. https://doi.org/10.1037/h0029439.
9. Aaron T. Beck, Robert A. Steer, and Gregory K. Brown. (1996). BDI-II Beck Depression Inventory-II. https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Personality-%26-Biopsychosocial/Beck-Depression-Inventory-II/p/100000159.
10. Aron, M., Beck, T. (1993). BHS Beck Hopelessness Scale. https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Personality-%26-Biopsychosocial/Beck-Hopelessness-Scale/p/100000105
11. Osman, A., Bagge. C. L., Guitierrez, P. M., Konick, L. C., Kooper, B. A., & Brrios, F. X. (2001). Suicide behaviour questionaire-revised. *Assessment, 443–454*(5), 4–6.

12. Recent, L. (2008). Columbia-suicide severity rating scale. 1–2. [Online]. Available http://www.integration.samhsa.gov/clinical-practice/Columbia_Suicide_Severity_Rating_Scale.pdf.
13. Evidence The Columbia Lighthouse Project. https://cssrs.columbia.edu/the-columbia-scale-c-ssrs/evidence/?fbclid=IwAR1c-ehlx-h1uhLLO48Z3L8Q-cOTOvJ2GqN2fM4xIPZPWusBlDEPIEbKGEA. Accessed Junuary 18, 2020.
14. The Columbia–Suicide severity rating scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3893686/?fbclid=IwAR3JEEA25peQp6OV3q-3pDF3i6kznJ2lnMX_l31uGWtspeciHgxg8QM-0Ki. Accessed Junuary 18, 2020.
15. Nilsson, C. M. E., Suryawanshi, S., Gassmann-mayer, C., XMcsorley, C., & Jiang, K. (2013). Columbia–Suicide severity rating scale scoring and data analysis guide. 1–13.
16. Subscale, I., et al. (2020). Clinical triage guidelines using the c-Ssrs, [Online]. Available http://cssrs.columbia.edu/wp-content/uploads/C-SSRStriageexampleguidelines.pdf. Accessed Junuary 18, 2020.

Detection of Antisocial Personality Based on Social Media Data

Shakil Mahmud, Masud Rana, Fahim Rubaiyat Zahir, and Mohammad Rezwanul Huq

Abstract Classifying antisocial personality characteristics is required to make a kind and tolerant society. Whatever people write and share in social media like Facebook subconsciously reveal his or her personality. At present, instead of psychologists, machine learning and data mining approach can make sound predictions of antisocial personality, also known as the dark triad. In this research, we surveyed a selected number of Facebook users based on the SD3 model. In this work, Random Forest (RF), Support Vector Machine (SVM), and Naïve Bayes (NB) algorithms have been used to classify dark triads such as Machiavellianism, narcissism, and psychopathy. Naïve Bayes provides better performance than Random Forest and SVM. However, our study also shows that psychopathy and narcissism are harder to detect than Machiavellianism.

Keywords Personality · Facebook · Dark triad · SD3 · SVM · RFC

1 Introduction

Successful classification of the antisocial personality could cause considerable changes to develop a good society. It is one of the most challenging tasks to accomplish. Social media platforms are currently the most popular place for internet users, offering them to communicate and share information with the whole world, and this online space offering social researchers an excellent opportunity to understand online

S. Mahmud (✉) · M. Rana · F. R. Zahir · M. R. Huq
East West University, Aftabnagar, Dhaka 1212, Bangladesh
e-mail: shakilmahmud13@gmail.com

M. Rana
e-mail: masudewu2016@gmail.com

F. R. Zahir
e-mail: fahimrz306@gmail.com

M. R. Huq
e-mail: mrhuq@ewubd.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_65

behavior. The most popular communication outlets in social networking are Facebook, Instagram, and Twitter [1] (2498 million, 1000 million, 386 million people, respectively), where the users find each other and interact, feel a part of a community and share their opinions. A total of 3.8 billion active users are engaged in social media, which is almost 49% of the overall population. And there is 8.7% or 304 million annual growth of the total social media user [2]. So social media is the biggest platform of data for the researcher, and this data is very widely used to predict the personality of a human. Using dishonest means that antisocial people usually believe that they are superior to others, and expect attention from others, although they excuse their misconduct. Still, somebody looks typical, charming, and they have lacking empathy and pathological lying characteristics. These characteristics of human behavior are called dark triad by psychology researchers, and they divided the dark triad into three parts: narcissism, psychopathy, Machiavellianism [3]. Several excellent studies had carried out to find an antisocial personality (dark triad) based on the SD3 method [4]. Facebook information has been used for detecting antisocial personalities in this research. Via Facebook, many characteristics of the dark triad naturally make themselves expressed. For example, narcissism is related to enhanced self-views, self-promotion, and craving for admiration, which can be reflected by selfie posting, always seeking attention, and projecting a positive view of the self [5]. A characteristic of psychopathy is antisocial and impulsive behavior expressed by an intensified usage of swear words. Infusion or manipulative language could be related to Machiavellianism [5].

2 Related Work

Social media data is distinctly the largest, most vibrant, and most compelling evidence for determining human behavior, bringing new opportunities to detect antisocial personality (dark triad). Data mining and machine learning techniques perform a significant role in finding the dark personality of humans based on their activity in social media. Jones et al. [6] represented major research demonstrating that SD3 offers solid, precise, and truly psychopathic, narcissistic, and Machiavellian measures. The study described by Wald et al. [4] showed the statistically significant associations between the big five characteristics and the dark triads. The big five characteristics are openness, extraversion, neuroticism, conscientiousness, and agreeableness. Preotiuc-Pietro et al. [5] mapped different characteristics to psychological theory and studied current forms of social media usage. Narcissism is linked to positive and somewhat banal expression. Psychopathy shows a very distinct trend of online activities, with most clusters of associated expression, from narcissism and Machiavellianism. The contents of violent, hostile posts relate to psychopathy and are distinguished by harmful impulses, whereas Machiavellianism indicates the fewest interactions of the three characteristics. Garcia et al. [7] concluded that psychopathy, narcissism, and Machiavellianism were significantly positively related and had satisfactory accuracy in the test. They also showed that narcissism was also associated

with the number of Facebook friends, whereas Machiavellianism did not depend on spending time on Facebook. Research has shown by Stiff [8] that the support of gossip fosters the relationship between Machiavellianism and Facebook surveillance.

In contrast, it is fostered by the disregard for uncertainty for psycho-path. Ryan et al. [9] tried to investigate who uses Facebook and the relationship between the big five, being shy, selfishness, being alone, and Facebook usage. And the results revealed that Facebook users are more narcissistic and extraverted. Carpenter [10] indicated that narcissism would be combined with antisocial attitudes.

3 Methodology

In this section, first, we discuss data collection procedures followed by the data preprocessing and feature extraction steps. Later, we explain the machine learning models we have used to obtain our results.

3.1 Data Collection

One hundred targeted Facebook users have initially taken part in this research, where 66 participants are males, and 34 participants are females. The age of the participants is between 18 and 55, and the average age is 23.6 years. The user dataset in this study has been divided into two parts—textual data and categorical data. Both types of data have been collected from the same set of users and then merged.

Firstly, the participants were requested to fulfill a Google survey form where Short Dark Triad (SD3) [3, 6] questionnaire has given for self-reported ratings (strongly disagree, disagree, neither agree nor disagree, agree, strongly agree) to measure narcissism, Machiavellianism and psychopathy. Finally, they have been asked to give their latest fifty Facebook status. But 30 participants have not offered their Facebook status, i.e., textual data. Therefore, 70 participant's data has been included in the dataset.

3.2 Data Preprocessing and Feature Extraction

Since the accuracy of a model is highly dependent on data, the collected data must be pre-processed. Facebook data contains many unnecessary characters. So, HTML tags, links, date, extra whitespaces, special characters, numbers, stop words, emoji, punctuation, and all characters except ASCII characters have been removed by using a regular expression. All texts have been turned into lowercase, and finally, stemming is applied by using NLTK Snowball Stemmer. Stemming is reducing inflected words to their word stem. Since text data cannot directly be fed into the algorithm, it

converted into a vector (sparse matrix of TF-IDF value) of numerical value using TfidfVectorizer [11]. Some features have been automatically removed while TF-IDF [12] transformation.

3.3 Proposed Model

The classification approach predicting the dark personality traits classifies a user's post to 'true' and 'false' class for each trait. The mean value of each trait calculated through the SD3 model is considered as the cut-off point for classifying. The following algorithms are used for classification. RF [13], SVM [14] and NB [15].

RF algorithm is an ensemble method that combines the result of decision trees. SVM is an algorithm that maximizes the boundary in the hyperplane between classes. Three types of kernels are used for classification—linear, polynomial, and radial. The above algorithms have been used for prediction. The dataset has been divided into 80% and 20% ratio for training and testing purposes. All text data have been tokenized and converted to vector by using TidfVectorizer [11] Scikit-learn library is used to implement TF-IDF and also RF, SVM, and NB algorithms.

Table 1 shows the tuned value of the parameters of the RF model. Here n_estimators is the number of decision trees in an RF, min_samples_splitis, and min_samples_leaf is the lowest number of samples required to split an inner vertex and to build a leaf vertex, respectively. Max_features is, think about the number of features when finding the best split (sqrt means the square root of the total number of features). The Bootstrap is a very effective sampling method.

Table 2 shows the tuned value of the parameters for all three kernels of the SVM. The parameter '*c*' is a regularization parameter, 'degree' is being used as the degree of the polynomial kernel, and gamma is the coefficient of the kernel.

Table 1 Model parameters for RF classifier

Parameters	n_estimators	min_samples_split	min_samples_leaf	max_features	bootstrap
Value	25	2	1	Sqrt	True

Table 2 Model parameters for support vector machine

Parameters	kernel	*c*	degree	gamma
Value	Linear, RBF, poly	1.0	9	auto

4 Performance Evaluation and Discussion

In this section, performance has been investigated based on applied algorithms and techniques.

Figures 1, 2, and 3 show the accuracy of the RF model for Machiavellianism, narcissism, and psychopathy is 0.9057, 0.5881, 0.6541, respectively. So, the RF model for Machiavellianism gives the best accuracy, which is 90% compared to narcissism and psychopathy.

Figure 1 illustrates that each of the SVM models for (linear, radial basis function (RBF), polynomial) Machiavellianism gives the same result that is about 91%. Figure 2 demonstrates for narcissism, the accuracy of the linear SVM model is (above 55%) a little bit lower than the RBF and polynomial model (above 57%). On the other hand, Figure 3 describes psychopathy, the accuracy of the linear SVM model is 68%, whereas the RBF and polynomial each model's accuracy are almost 68%. So, the SVM model is also the best classifier for Machiavellianism.

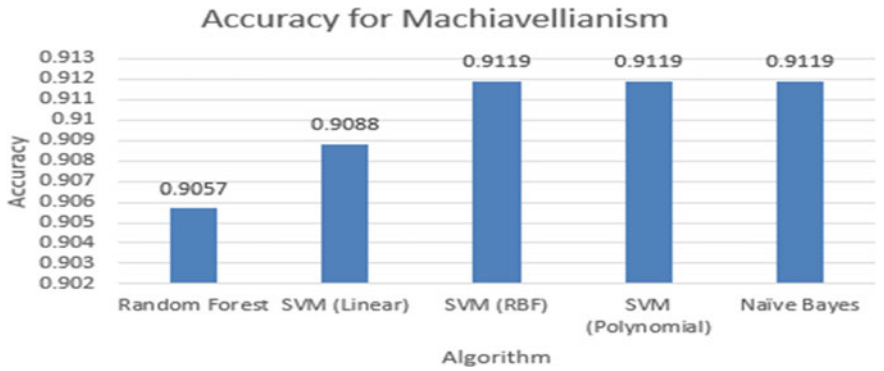


Fig. 1 Accuracy of different classifiers for Machiavellianism

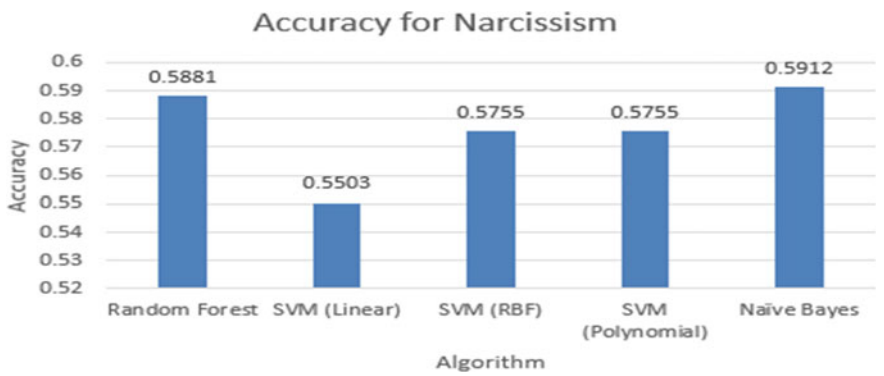


Fig. 2 Accuracy of different classifiers for narcissism

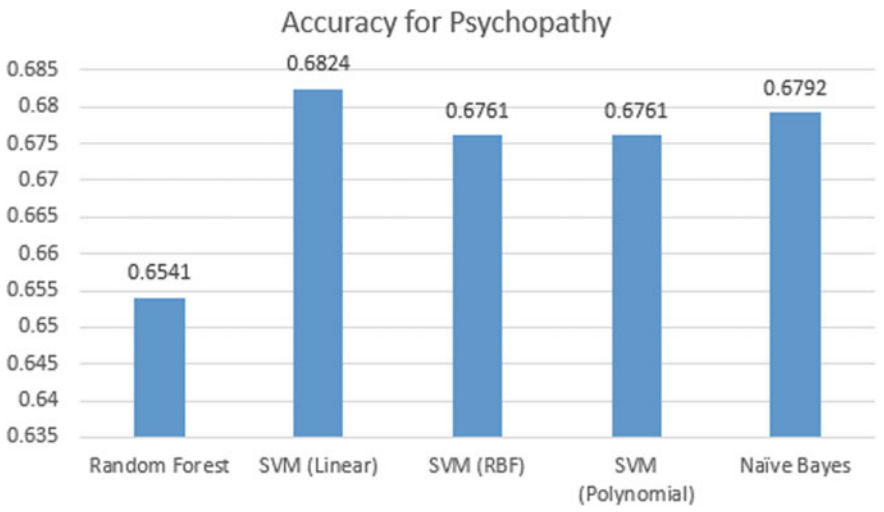


Fig. 3 Accuracy of different classifiers for psychopathy

Figures 1, 2, and 3 also represent that accuracy of Naïve Bayes for Machiavellianism, narcissism and psychopathy is about 91%, 60%, 68%, respectively. So, it is clear that for Machiavellianism, all of the algorithms give better accuracy where NB and SVM (linear, polynomial, RBF) (91%) provide 1% better accuracy than RF (90%).

Figure 4 shows the precision of the RF model for Machiavellianism, narcissism, psychopathy, which is 0.9114, 0.5278, 0.3333, respectively, whereas recall for Machiavellianism, narcissism, psychopathy is 0.996, 0.191, 0.077, respectively. And the F1_score is 0.9505, 0.3615, 0.1129, respectively, for the following three traits.

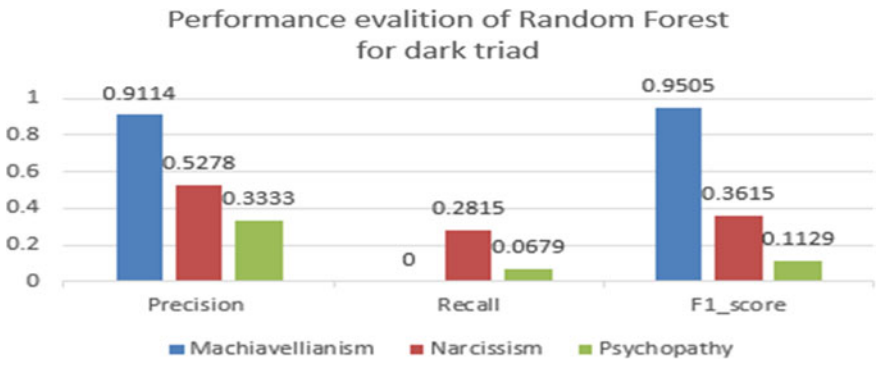


Fig. 4 Precision, recall, and F1_score for Machiavellianism, narcissism, psychopathy using RF algorithm

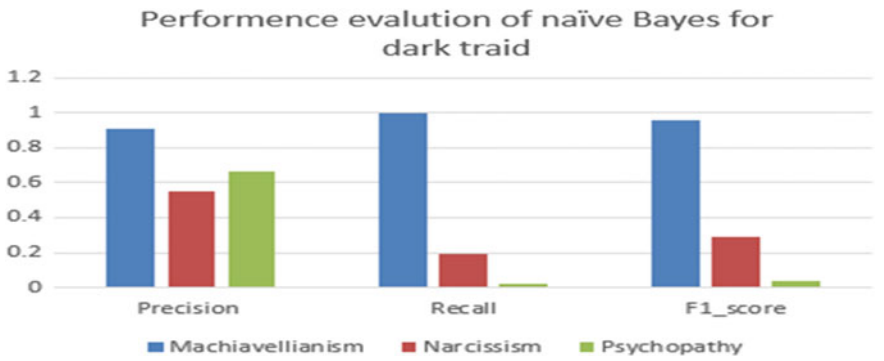


Fig. 5 Measuring precision, recall and F1 score for Machiavellianism, narcissism, psychopathy using Naïve Bayes algorithm

Table 3 The result variance of Machiavellianism, narcissism, psychopathy

Dark personality traits	Data variance
Machiavellianism	0.255
Narcissism	0.274
Psychopathy	0.413

Figure 5 illustrates the precision of the NB model for Machiavellianism, narcissism, and psychopathy is 0.9119, 0.5532, 0.667, respectively, whereas recall for Machiavellianism, narcissism, psychopathy is 1.0, 0.192, 0.0194, respectively. And the F1 score is 0.9539, 0.2857, 0.0377, respectively, for the following three traits. Comparing both RF and NB, it can be said that NB gives a little better precision, recall, F1 score than RF.

After analyzing the result, it can be concluded that NB performs better than RF and SVM. However, NB and SVM predict the same level of accuracy for Machiavellianism, and all of the classifiers can make better predictions for Machiavellianism than narcissism and psychopathy. Table 3 shows the data variance of Machiavellianism, narcissism, and psychopathy, which is 0.255, 274, and 0.413, respectively, calculated from the survey. Table 3 shows that the data variance of Machiavellianism is lower than narcissism and psychopathy. The less variation for Machiavellianism justifies why all classifiers make a better prediction for this criterion.

5 Conclusion and Future Work

An acceptable means of detecting antisocial personality traits are required to develop a sound and tolerant society. RF, Support Vector Machine (SVM) (linear, polynomial, RBF kernel), Naive Bayes (NB) are applied to classify the antisocial personality based on social media data. The study ensures that NB can produce a better classification

result than the SVM (linear, polynomial, RBF kernel) and RF. It is clearly shown that both RF, SVM, and NB models can make a highly accurate classification result for Machiavellianism compared to narcissism and psychopathy.

This study could not achieve the expected accuracy rate for psychopathy and narcissism classification because of the considerable variance in collected data. In the dataset, there may be a lack of variety of data because the survey result shows that the participant's age range is almost the same, all of the participants are Bangladeshi, and less than a hundred participants participated in the survey. In the future, using a much larger dataset, the accuracy of the antisocial personality classification scheme can be further elevated by deep learning-based approaches such as LSTM, GRU, CNN.

References

1. Dustin, W. S. (2020). Social Media Statistics 2020: Top Networks By the Numbers • Dustin Stout. https://dustinstout.com/social-media-statistics/. Accessed June 6, 2020.
2. Kemp, S. (2020). Digital 2020: 3.8 billion people use social media—We are social. https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media. Accessed 6 Jun 2020.
3. Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality*. https://doi.org/10.1016/S0092-6566(02)00505-6.
4. Wald, R., Khoshgoftaar, T., & Sumner, C. (2012). Machine prediction of personality from Facebook profiles. In *Proceedings of the 2012 IEEE 13th International Conference on Information Reuse and Integration, IRI* 2012.
5. Preoţiuc-Pietro, D., Carpenter, J., Giorgi, S., & Ungar, L. (2016). Studying the Dark Triad of personality through twitter behavior. In *Proceedings of International Conference on Information and Knowledge Management*.
6. Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3): A brief measure of dark personality traits. *Assessment*. https://doi.org/10.1177/1073191113514105.
7. Garcia, D., & Sikström, S. (2014). The dark side of Facebook: Semantic representations of status updates predict the Dark Triad of personality. *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2013.10.001.
8. Stiff, C. (2019). The dark triad and Facebook surveillance: How Machiavellianism, psychopathy, but not narcissism predict using Facebook to spy on others. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2018.12.044.
9. Ryan, T., & Xenos, S. (2011). Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. In: *Computers in human behavior.*
10. Carpenter, C. J. (2012). Narcissism on Facebook: Self-promotional and antisocial behavior. *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2011.11.011.
11. sklearn.feature_extraction.text.TfidfVectorizer—scikit-learn 0.23.1 documentation. https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html. Accessed June 6, 2020.
12. Scott, W. (2019). TF-IDF from scratch in python on real world dataset. https://towardsdatascience.com/tf-idf-for-document-ranking-from-scratch-in-python-on-real-world-dataset-796d339a4089. Accessed June 6, 2020.
13. Synced. (2017). How Random Forest Algorithm works in machine learning. https://medium.com/@Synced/how-random-forest-algorithm-works-in-machine-learning-3c0fe15b6674. Accessed June 6, 2020.

14. Support Vector Machines—scikit-learn 0.23.1 documentation. https://scikit-learn.org/stable/modules/svm.html. Accessed June 6, 2020.
15. Document Classification Using Multinomial Naive Bayes Classifier. https://www.3pillarglobal.com/insights/document-classification-using-multinomial-naive-bayes-classifier. Accessed June 6, 2020.

Devising a Novel Means of Introducing Collaborative Decision-Making to an Automated Water Bottling Plant to Study the Impact of Positive Drift

J. Coetzer, R. B. Kuriakose, and H. J. Vermaak

Abstract Cycle time refers to the amount of time allotted to accomplish a certain process in an assembly process. The deviation from the optimal cycle time is technically referred to as drift. Drift can be negative or positive. Positive drift represents the time span in which an assembly line exceeds the prescribed cycle time. While previous studies have shown that an automated assembly line can reduce the positive drift considerably, this paper looks to devise a novel means of introducing a collaborative decision-making framework involving humans as well as machines to reduce the impact of positive drift. This is done using the case study of a water bottling plant. The paper initially showcases the automated operation of the water bottling plant. Secondly, the paper examines where collaborative decision-making can be introduced in the plant. Thirdly, a set of preliminary results which point to the effectiveness of the collaborative approach is described. The paper is concluded by looking at how the approach can be expanded to create a generic framework for similar plants in the future.

Keywords Collaborative decision-making · Collaborative approach · Positive drift

1 Introduction

Collaborative decision-making is used to facilitate efficient science-based decision-making and in the context of this research, can be defined as an approach whereby humans and machines work together on a specific task to reach a common goal [1]. The activity of collaboration between humans and machines is similar to how

J. Coetzer (✉) · R. B. Kuriakose · H. J. Vermaak
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

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_66

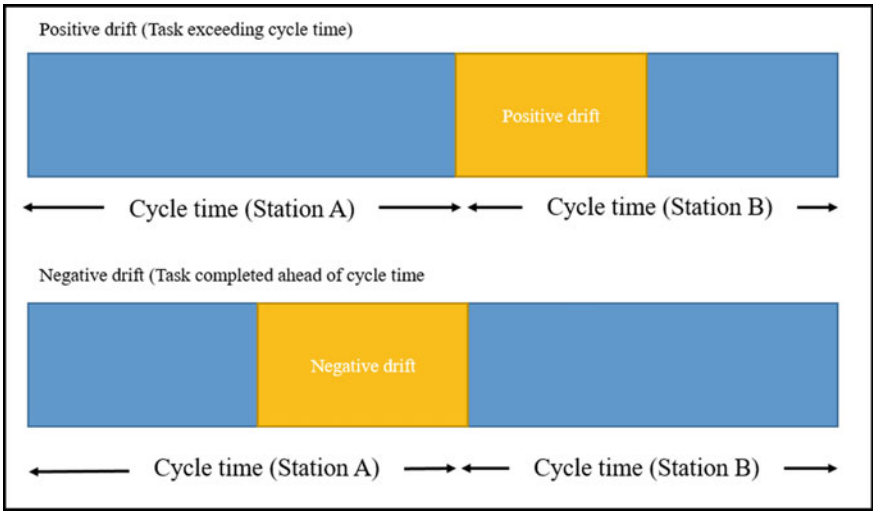


Fig. 1 Positive and negative drift

humans collaborate for solving problems whereby both parties involved bring their own special skills to the process [2].

The main advantage of implementing collaborative decision-making in the production process is that the strengths of both humans and machines can be utilized. This will result in the solving of more complex problems than when working in isolation [2].

One of the complex problems that a production process endures is that of positive drift. Drift refers to the deviation from an optimal cycle time [3]. Drift can be positive or negative (see Fig. 1). As seen from Fig. 1, positive drift occurs when work assigned to a station is not completed within the allotted cycle time (cycle time of Station A) and extends into the cycle time of the next workstation (cycle time of Station B).

The effect of positive drift can be reduced by automating the production process. A study that was done preceding [4] to the current one has garnered enough evidence in support of this fact. This research was conducted on an automated water bottling plant [5, 6] which will also be used as a case study in this research.

This paper explores how collaborative decision-making can be introduced to the automated water bottling plant. A comparative study will then be done to measure the impact that both the approaches have on positive drift in the water bottling plant.

The paper is structured such that it initially outlines the automated water bottling plant which will be used as a case study in this article. Secondly, it will explain in detail collaborative decision-making and how it can be introduced to the water bottling plant. Thirdly, the paper will discuss preliminary results from the research and finally conclude on how the work can be expanded.

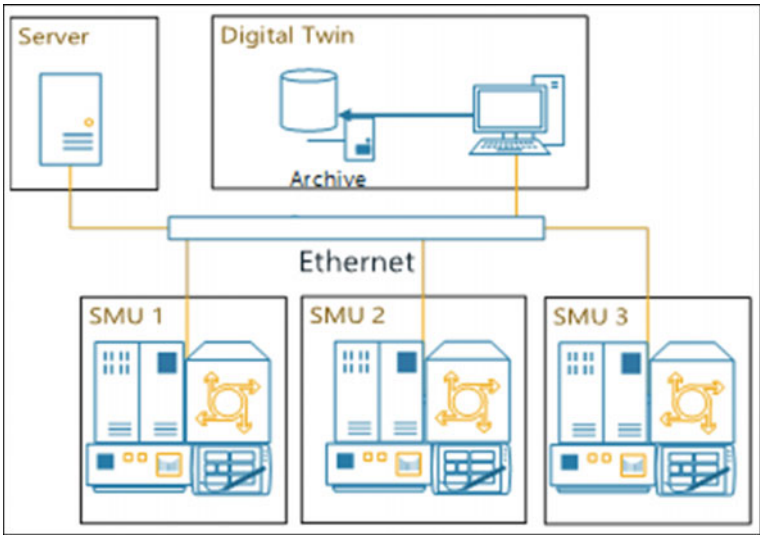


Fig. 2 Digital twin and smart manufacturing unit

2 The Water Bottling Plant

The automated bottling plant is split into three major subsections. Subsection A contains the water source, while subsection B is the bottle storage, and subsection C is the water filling unit. The aim of the plant is to execute customer orders for 500 ml and 750 ml bottles through a cloud server.

The plant was modelled in Simulink [6] and built as a digital twin [7] to analyse its operation before three smart manufacturing units (SMUs) were constructed using Siemens S7 programmable logic controllers (PLCs) to commence the production process. A diagram depicting the SMUs connected to the digital twin is shown in Fig. 2.

3 Collaborative Decision-Making

Collaborative decision-making is used to facilitate efficient science-based decision-making and in the context of this research, can be defined as an approach, whereby humans and machines work together on a specific task to reach a common goal [1]. The activity of collaboration between humans and machines is similar to how humans collaborate for solving problems. Both parties involved bring their own special skills to the process, similar to human collaboration [8]. By implementing collaborative decision-making in the production process, advantage can be taken from the strengths

of both humans and machines resulting in the solving of more complex problems and finding better solutions to problems than when either would be when working in isolation [9].

3.1 Introducing Collaborative Decision-Making to the Automated Water Bottling Plant

In order to test an automated system for human–machine collaboration, a human–machine interface (HMI) will be created and connected to the SMUs. This will aid in enabling collaborative decision-making between the human operator and the system. A supervisory control and data acquisition system (SCADA) will be developed and used to provide the different configuration options for the water bottling plant.

A SCADA is an integrated control system that utilizes computer functions, graphical user interfaces and networked data communications [10] for the control, data acquisition and monitoring of different aspects of industrial plants and manufacturing environments [8, 11]. The SCADA system has four functions: data acquisition, network communication, data presentation, data control and monitoring [11].

SCADA systems are a combination of many subsystems including sensors, remote terminal units (RTUs) and input and output devices connected to programmable logic controllers (PLC) which is used to regulate the process or machinery [10]. Data collected from the subsystems are transmitted to the central SCADA unit which has its own human–machine interface (HMI) connecting users directly with the system enabling data acquisition, monitoring and data controlling [11].

Human–machine interaction refers to the communication and interaction between a human and a machine by means of a user interface [8]. A human–machine interface (HMI) is a system enabling interaction between humans and machines and is essentially a subset of a SCADA as shown in Fig. 3. The HMI increases interaction between the user and the machine through computer screens or touch screens while aiming to meet the user needs for system information [12].

The HMI of the SCADA displays processes that happen throughout the whole system by utilizing a graphical user interface (GUI) allowing the user to visually see and comprehend the functioning and status of the system [12] as shown in Fig. 3. In this project, a SCADA with an HMI will be connected to an automated water bottling plant. The aim of which is to determine if a human–machine collaboration can result in reducing positive drift as opposed to running the system automatically.

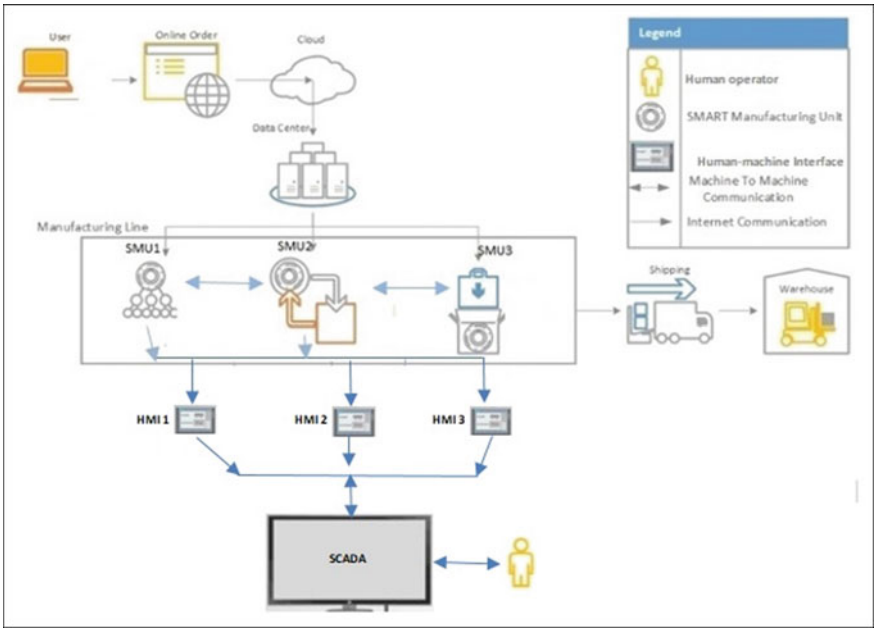


Fig. 3 Layout of the water bottling plant with SCADA and HMI

3.2 A Model of the SCADA and HMI for the Automated Water Bottling Plant

The layout of the water bottling plant with SCADA and HMI is depicted in Fig. 3 and shows the three SMUs which will be used for filling the water bottles. The SMUs were constructed using PLCs and have level sensors and proximity sensors which function to gather data and send it to the SCADA.

There is an HMI connected to each of the SMUs on the water bottling plant. The HMIs are in turn connected to the SCADA which monitor and control elements connected to it. Based on information projected on the SCADA, the HMI can be used by the human operators to directly control the processes in the plant.

The HMI aims to increase the interaction between the system and the operator and is designed with an intuitive touch-based input screen for effective and efficient communication by the operator with the system. The graphical user interface (GUI) on the HMI will display the system information graphically by means of user-friendly buttons, levers or sliders with which the human can interact when intervention is needed.

The SCADA is programmed to create an alert when certain constraints are met such as when the water reaches a certain level or when there are no more bottles in storage to be filled and capped. The HMI can then be used to control the speed of production. The human can take a decision to adjust the flow rate depending on the

availability of water or to leave the speed unchanged and whether the speed of the bottle capping should be adjusted or halted. In addition to the HMI creating alerts, the SCADA will supply the human with an overview of the status of the system to enable the human to take informative decisions for continuing or halting of the automated process.

4 Preliminary Results and Discussions

This section aims to compare the results from the first HMI (HMI-1) of the water bottling plant. HMI-1, as explained in Sect. 3, when connected to the water storage SMU (SMU-1). A flow sensor as well as a level sensor is connected from the SMU to the SCADA to monitor the flow of water from the tank and the level of water in the tank.

The automated operation of the tank is optimized such that alerts are created when the water level reaches 50 and 25%. At 25% capacity, the production process is halted, and water in the tank is replenished to at least 75% capacity. This is shown in a decision tree diagram shown in Fig. 4.

It is clear from Fig. 4 that there is delay at 441 min when the water level in tank is at 25%. It is also visible that after the tank has reached 75% capacity, there is variation in the flow rate of water from the tank. The process is halted when water level is zero. This is in line with the optimized operation of the plant for a given set of customer inputs.

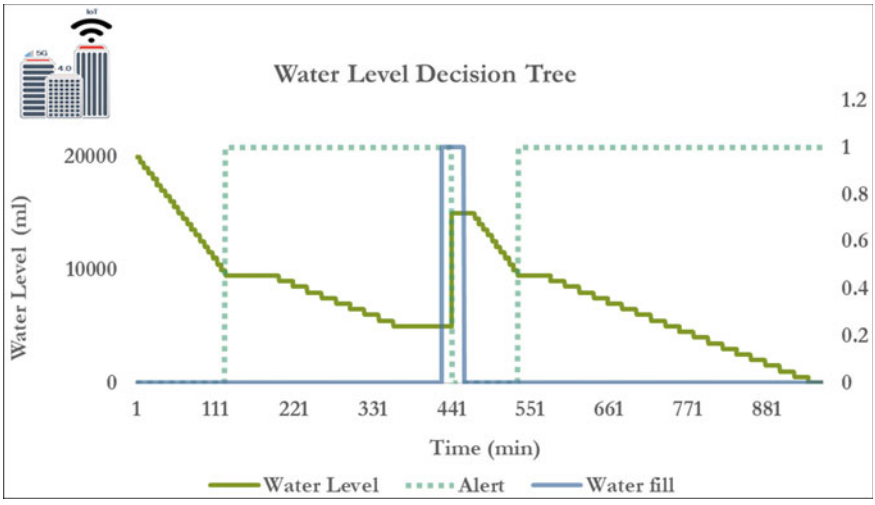


Fig. 4 Decision tree diagram showing the operation of the water plant in automated mode

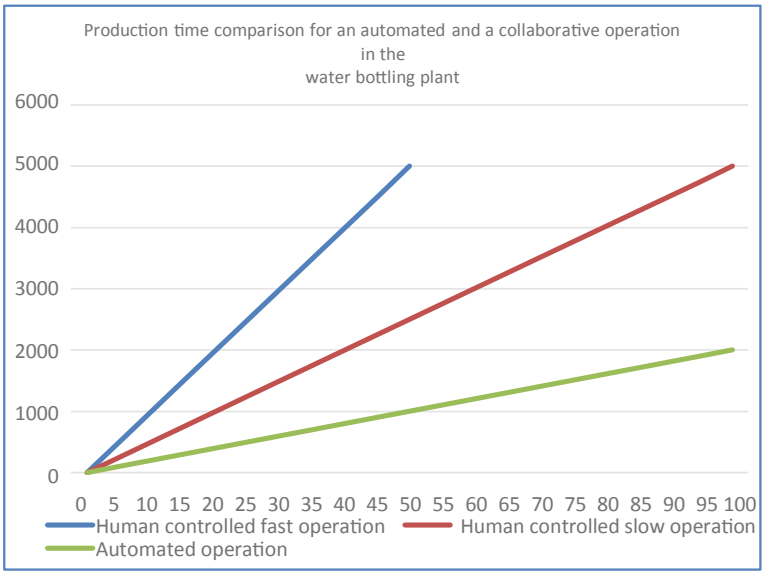


Fig. 5 Production time comparison for an automated and collaborative operation of water bottling plant

In the experiment conducted in this section, the alerts created for the automatic operation have been removed and replaced with HMI which will be used to vary the flow rate of water based on information from the SCADA by a human being. It is of importance to note here that the halting of the process when water level is zero is still controlled by the machine and the human intervention is only introduced at HMI-1, thereby making the process a collaborative decision. The results are depicted in Fig. 5.

The test is designed to compare the production time for an automated and collaborative operation for bottling 5000 bottles. As seen from Fig 5, the automated operation has only reached 2000 bottles in 100 min. The collaborative approach is split into a fast as well as a slow approach.

The fast approach will drain the tank faster, but using the SCADA, the level of water can be monitored, and the flow rate can be controlled using the HMI. In the slow approach, the water will be drained at a slower rate than the fast operation, but faster than the automated approach. In both the fast and slow collaborative controlled approaches, the production process is completed faster than the automated approach, and the process was never halted.

This proves that for this specific section of the plant, positive drift can be reduced considerably by introducing a collaborative decision-making. However, this does not prove that a human intervention will reduce the positive drift of the entire plant. It could well be that in a specific section of the plant, the automated approach is much faster than the human intervention. Further study needs to be done in this respect and will form part of the future works.

5 Conclusion

The purpose of this research was to determine the impact of collaborative decision-making on positive drift in an automated production environment. For this purpose, a case study using an automated water bottling plant was chosen. The automated plant was modelled in Simulink and constructed using S7 Siemens PLCs.

The automated operation is based on an optimized algorithm which is determined by the input and the constraints such as the level of water in the storage, number of bottles available in the storage and the flow rate of water. A collaborative approach was brought to SMU-1, where the level and flow rate of water are constraints.

An HMI connected to SMU-1 and the SCADA was used to control the flow rate of water, while the level of water was controlled by the machine. This meant that the production process would be halted if the water level in the tank was zero. Otherwise, the production process would be controlled by human.

The SCADA gives information about the level of water in the tank, and based on that data, the human being was able to control the flow rate of the water. In order to test the effectiveness of collaborative decision-making, two scenarios were tested here, a fast and a slow approach. Both these approaches were compared to the automated approach.

The tests showed that, in this specific case, the collaborative controlled approach was faster that the automated approach. However, this only applies to one section of the plant and not the entire functioning of the plant. As future work, the research will measure the impact of introducing a collaborative controlled approach in different sections of the operation of the plant.

The aim is to create a framework which details how best to combine the perceptual abilities of human operators and automated processes to function on a collaborative manner to reduce the impact of positive drift and increase efficiency in the production process.

References

1. Ulibarri, N. (2018, February). Collaborative model development increases trust in and use of scientific information in environmental decision-making. *Environmental Science & Policy, 82*, 136–142. https://doi.org/10.1016/j.envsci.2018.01.022.
2. Allen, J., & Ferguson, G. (2002). Human-machine collaborative planning. In *International NASA Workshop on Planning* (pp. 1–10). Retrieved from https://www.cs.rochester.edu/research/cisd/pubs/2002/allen-ferguson-nasa2002.pdf.
3. Hu, S. J., et al. (2011). CIRP annals—Manufacturing technology assembly system design and operations for product variety. 60, 715–733. https://doi.org/10.1016/j.cirp.2011.05.004.
4. Kuriakose, R. B., & Vermaak, H. J. (2019). Optimization of a real time web enabled mixed model stochastic assembly line to reduce production time. *Communications in Computer and Information Science, 1075*(1), 39–51.
5. Kuriakosem R. B., & Vermaak, H. J. (2019). Customized mixed model stochastic assembly line modelling using simulink. 2–7. https://doi.org/10.5013/ijssst.a.20.s1.06.

6. Kuriakose, R. B., & Vermaak, H. J. (2020). Designing a Simulink model for a mixed model stochastic assembly line: A case study using a water bottling plant. *Journal of Discrete Mathematical Sciences and Cryptography, 23*(2), 329–336.
7. Kuriakose, R. B., & Vermaak, H. J. (2019). Optimization of a customized mixed model assembly using MATLAB/Simulink. *Journal of Physics: Conference Series*, *1201*(1).
8. Coetzer, J., Kuriakose, R. B., & Vermaak, H. J. (2019). Collaborative decision making for human technology interaction—A case study using an automated water bottling plant. In *International Conference on Electronics Representation and Algorithm.*
9. Noordin, N. A., Bititci, U. S., & Van Der Meer, R. (2012). Review on collaborative decision making in supply chain: The relationship between e-collaboration technology and development of inter-organizational trust. In *IFIP International Conference on Advances in Production Management Systems* (Vol. 384 AICT, pp. 326–341). https://doi.org/10.1007/978-3-642-33980-6_37.
10. Chhabra, A. (2019). The Relation between IIoT, SCADA and HMI Explained. *Industrial Software.* [Online]. Available https://blog.se.com/industrial-software/2019/05/28/the-relation-between-iiot-scada-and-hmi-explained/.
11. Setiawan, A., Sugeng, Koesoema, K. I., Bakhri, S., & Aditya, J. (2019). The SCADA system using PLC and HMI to improve the effectiveness and efficiency of production processes. In *IOP Conference Series: Materials Science and Engineering* (Vol. 550). Institute of Physics Publishing. https://doi.org/10.1088/1757-899X/550/1/012008.
12. Fitriani, S., & Sofyan, Y. (2020). Simulator human machine interface (HMI) using visual basic on the SCADA system simulator human machine interface (HMI) using visual basic on the SCADA system. In *IOP Conference Series: Materials Science and Engineering* (Vol. 830). https://doi.org/10.1088/1757-899x/830/3/032016.

Cluster-Based Ensemble Using Distributed Clustering Approach for Large Categorical Data

Govind Pole and Pradeepini Gera

Abstract The modern data analytics approaches suffer through the various challenges when dealing with large datasets, and thus, there is qualitative and quantitative degradation in the decision-making process. There is a need of an efficient analytical system that can handle on large and diverse datasets with increased accuracy in data analysis. Here, a prototype model of a cluster-based ensemble system is proposed to run on commodity machines using hierarchical approach. The commodity machines create the cluster-based ensemble in a distributed environment using a popular data mining algorithm. The experimentation proves that hierarchical approach helps to decrease time for data mining process, in turn boosts the speed of decision-making and the cluster-based ensemble improved the accuracy.

Keywords Distributed clustering · Categorical data · Ensemble learning · Distributed analysis · K-mode clustering · Commodity computing

1 Introduction

The big data analysis governs extremely large volumes, high varieties, different formats with growing data sets size of multiple types like numeric, non-numerical/categorical data and mixed data, etc. [1]. The analysis of such data, which is qualitative and non-numerical data, is an active research field in statistics, machine learning, data mining, artificial intelligence, etc. Toady, large amount of research is going for understanding and analyzing such descriptive, non-numeric categorical data. The mean of objects in cluster of categorical data cannot be calculated like numeric data, because they lack intuitive arithmetic properties.

G. Pole (✉) · P. Gera
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India
e-mail: govindpole@gmail.com

P. Gera
e-mail: pradeepini.gera@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_67

Many algorithms have been implemented for a various clustering tasks, still they suffer from the limitations because of unsupervised learning [2]. They are model selection and proper grouping. The first one is searching a solution that addresses the deep-down number of clusters in original data set, while the second one concentrates to assembly coherent elements together to form a cluster. Here, it focuses on the proper grouping and is intended to boost the clustering analysis performance using ensemble techniques. Due to the inherent ability of exploiting advantages of combining the numerous clustering solutions, cluster-based ensemble mostly improves upon the single clustering algorithm in accuracy, novelty and reliability. Cluster-based ensembles work in two steps. Initially, multiple partitions are generated as ensemble members with different initialization for same algorithm or using different clustering algorithms. Later, a specific consensus function combines the multiple partitions into the consensus partition [3]. There are consensus functions developed such as hyper-graphic partitioning [4], co-association-based consensus function [5–7], voting-based consensus function [8, 9], etc. Different strategies to generate the input partitions also have been developed like using multiple clustering algorithms to generate the input partitions [4], using a single algorithm for clustering by using different initialization and different parameters [10], using single algorithm for clustering on different features extracted from the original data [11], and applying a single clustering algorithm to the subset sampled from the original data [4].

This paper proposes a cluster-based ensemble approach for clustering the categorical data. It is outlined as follows. Section 1: Introduction of cluster-based ensemble. Section 2: The K-mode clustering algorithm and its variant for categorical data. Section 3: Cluster-based ensemble using distributed clustering approach. Section 4: Result discussion. Section 5: Conclusion and future Scope.

2 The K-Mode Clustering Algorithm and Its Variant

The K-mode clustering technique is immerged to overcome the limitations of a well-known and prominent clustering algorithm named K-means clustering which is built on idea of distance measure. The limitation of this algorithm is that it is useful only for numeric dataset. To tackle this problem, the k-modes algorithm is proposed which handles categorical dataset effectively. The remarkable advantage of the K-modes clustering is that it is faster to K-means clustering, as it needs less number of iteration than the counterpart [12, 13].

There are varieties of clustering algorithms available that are K-means, K-medoids or mean-shift, k-prototype but none of the algorithm returns K meaningful modes. The K-mode algorithm has the combination of the ideas of density and cluster assignment. The first idea is, k-mode uses measure of matching dissimilarity for grouping of categorical variables, rather than similarity computation. Secondly, it uses mode of objects instead of means of objects for clusters as it is difficult to find means for categorical or non-numeric values. Finally, it uses object appearance frequency to update modes to minimize the clustering complexity [13].

There exist different k-mode amendments causing a family of k-mode algorithms.

2.1 Clustering Using Similarity Weight and Filter Method

Here, the ensemble approach involving combination multiple executions of different clustering algorithm on same set of mixed data is used for the optimization of the mining task. Thus, the cluster ensemble produces a common optimized cluster of the given dataset [15]. In this cluster ensemble approach, quantitative and qualitative data is handled separately and then the merged data processed as a qualitative data. Several algorithmic techniques are popular for such non-numeric like "K-m algorithm" [13], "K-prototype algorithm" [16], "Rock" and "squeezer algorithm" [17]. Following approach was used to archive this.

1. The original dataset is separated as numerical data and categorical data.
2. The best in class existing clustering algorithms is applied to both numeric and categorical/nominal data set.
3. Combine the output of above algorithms (used in Step 2).
4. Finally, treat the result as categorical data and apply squeezer algorithm to result of Step 3.

2.2 Clustering Using Weighting Uncommon Attribute Value Matches

The contemporary algorithms for clustering categorical data treat all attributes equally for similarity computation. Thus, the frequent and rare attributes matching have same weight for nearness computation. This work is an attempt to improve the clustering accuracy by giving more weightage to uncommon attributes matching in similarity calculations [18]. Here, "squeezer algorithm" [17] and the modified squeezer algorithm are compared for performance and later was found more accurate than the original one.

2.3 Haitian's Method for K-Modes Clustering

It is a based on idea of optimization heuristic which considers the very precise change to the clustering distance function before moving an element from one group to another. There is very less research on how to use this optimization heuristic on k-modes type algorithms. This work devises an initialization of above method based on heuristics and shows that its computational complexity on k-modes clustering is as good compared to "Lloyd's method" on k-modes clustering. This method overcomes the problem of a local optima [19].

3 Cluster-Based Ensemble Using Distributed Clustering Approach

The distributed processing has advantages of high-performance with less costly machines. So a distributed classifier system implementation is used for high performance in data mining process.

3.1 Cluster-Based Ensemble

Different algorithmic solutions are provided for same data mining tasks, so many times there is a dilemma that which algorithm or set of algorithms to be considered for solving the problems with optimum results [21]. Cluster-based ensemble is a process of optimizing clustering result by combining multiple execution of the different algorithm on a dataset to get single consolidated clustering. This approach does not look inside the features or algorithms that determined these partitioning. Several research has been carried out on ensemble-based distributed clustering, based on K-means, expected maximization (EM), and hierarchical clustering methods [2]. Here, we proposed a distributed hierarchical clustering algorithm for categorical data which requires multiple rounds of message passing. Here, we will cluster the categorical data through the different clustering algorithms (1) k-means one-hot encoding: first categorical data is converted to numeric data and then clustering k-means. (2) K-modes: It uses modes of categorical attributes rather than the means in k-means for clustering of categorical data.

3.2 The Design of Distributed Analysis System

It is a model designed based on the concepts of distributed systems, which consists of peer to peer connected nodes that performs analysis task. The nodes are called as analyzers. One of the nodes forms the network act as a coordinator, which coordinates the analysis task and also involved in data analysis. The data communication occurs in two phase, first distributing data and receiving analysis result.

The set of commodity machines or servers are associated with analyzer, which performs data analysis on these systems. The set of analyzers which are associated with one coordinator system and all coordinators are linked to each other. The updates on commodity machine, i.e., nodes are analyzed by associated analyzer. Then the analyzer systems forward the updates to associated coordinator, which in turn performs the merging and creates the final result [23].

3.3 Distributed Cluster Generation

Algorithm: The distributed cluster generation

(1) The dataset D consists of n samples ($D = \{d_1, d_2 \ldots, d_n\}$), and each sample di is represented by a feature vector ($d_i = \{d_{i1}, d_{i2}, \ldots, d_{i1}\}$) (where $i \in \{1, \ldots, n\}$ and m is the number of features).
2) The random subspace technique is used to generate a set of new datasets $\{D_1, D_2, \ldots, \text{DB}\}$ in the first step. The sample is chosen oneat a time until S_n samples are procured.
(3) Then, the different clustering algorithms like K-mean-ohe, K-mode and Rock are used to perform clusters in the subspaces, and generate a set of C_1, C_2, … CB.
(4) Finally, a consensus matrix is constructed using the set of different solutions and the end result is obtained.

3.4 The Cluster-Based Ensemble Selection

Algorithm: Cluster-based ensemble selection (CBES)
Input: The dataset D where ($D = \{d_1, d_2 \ldots, d_n\}$), and each sample di is represented by a feature vector ($d_i = \{d_{i1}, d_{i2}, \ldots, d_{i1}\}$) (where $i \in \{1, \ldots, n\}$ and m is the number of features)

Step 1: The random subspace technique is used to generate a set of new datasets $\{D_1, D_2, \ldots, \text{DB}\}$ in the first step. The sample is chosen one at a time until Sn samples are procured.

Step 2: Run consensus clustering across subsamples of data, clustering algorithms and cluster sizes.

Step 3: Computes the internal validity indices for a clustering result and corresponding compactness score.

Step 4: Combines results for multiple objects from Steps 2, 3 and outputs the consensus matrices.

Step 5: Evaluate algorithms based on the on internal/external validation indices calculated in Step 3.

Step 6: Best performing algorithms from Step 5 are selected as a member of ensemble.

Step 7: Then, the selected members set used are used to perform clustering in the subspaces, and generate a set of C_1, C_2, …, CB.

Step 8: Finally, a consensus clustering is performed based on clusters in Step 7 and the final cluster is obtained.

Table 1 Summary of dataset

Dataset	Instances	Attribute	Sources
Mushroom	8124	22	UCI
US Census Data	2,458,285	68	UCI

3.5 Datasets

Standard data set collections are used for Experiments. The USCensus1990 raw data set contains one percent sample 1990 census dataset [20]. There are 68 categorical attributes. Another dataset is mushroom dataset which described in terms of physical characteristics this data set is downloaded from UCI machine learning laboratory [22] (Table 1).

4 Results and Discussion

In the cluster-based ensemble system, we are more concern about processing large and high dimensional data for analysis in reasonable amount of time. To achieve this, we are using set of commodity machines and servers, to perform the complex and time consuming data mining task on datasets. Here, we will discuss our experimental results for above mentioned cluster-based ensemble system. To speed up for performance, cluster formation of datasets is carried out using distributed approach. We have taken the entire mushroom sample dataset with 8124 samples and only 450 samples from US census dataset. Less number of samples helped in ensemble experimentation trials and their graphical representations. In Table 2, we can see that the number as the number of clusters increases, the time required for clustering increases. So the when doing ensemble, we can delegate the tasks to different nodes and fetch only the cluster computations for selecting the ensemble member. Again the cluster parameters like compactness score within cluster difference help us to select the ensemble members. From Tables 3 and 4, we can say that the purity of cluster can play the significant role for selecting the cluster ensemble member. The consensus

Table 2 Cluster analysis for US census dataset

No.of samples	Clustering time in	No. of clusters	Cluster size	Within cluster difference	Compactness
450	23.64835 s	2	310, 190	4456, 2756	2409.353
450	28.67467 s	3	205, 111, 134	3260, 753, 1935	2406.859
450	34.77027 s	4	131, 108, 105, 107	1907, 728, 1475, 1530	2418.699
450	39.97434 s	5	144, 54, 141, 70, 42	966, 322, 1878, 939, 272	2405.657

Table 3 Cluster analysis for mushroom dataset

No. of samples	Clustering time	No. of clusters	Cluster size	Within cluster difference	Purity
K-means-ohe 8124	9.491989 s	2	4993, 3130	40157, 22522	0.895482
	10.32873 s	3	1296, 1845, 4982	4860, 9741, 40046	0.8946202
	10.8651 s	4	3246, 1850, 1296, 1731	26672, 9793, 4860, 7816	0.8940047
	11.32607 s	5	1056, 1731, 1852, 1296, 2188	540, 7821, 9815, 4860, 17837	0.8937585
K-modes 8124	10.64841 min	2	5023, 3100	40507, 22486	0.8000739
	12.19493 min	3	5689, 255, 1179	49520, 8116, 7183	0.6388034
	13.69575 min	4	2342, 3294, 1424, 1063	14,969, 24,922, 9350, 5875	0.691124
	20.70787 min	5	2630, 1006, 2084, 1335, 1068	15378, 5569, 12334, 6509, 5912	0.8618737

Table 4 Purity of cluster for mushroom dataset

Algorithm	No. of samples	Clustering time	Purity
K-means-ohe	8124	9.491989 s	0.895482
K-modes	8124	10.64841 min	0.8000739

clustering creates the cluster assignments across subsamples using different clustering algorithms HC, PAM, DIANA, k-means, for different k number of clusters. Here, US census data is clustered into three or four clusters, using 80% sampling rate with 5 times replication. Different internal indices like c-index, calinski_harabasz, Davies_Bouldin, pbm, etc are computed and summarized in Table 5. They are used comparison and analysis of clustering algorithms performance (Figs. 1 and 2).

5 Conclusion

The distributed clustering fetches the computations (cluster values) from one level to another, rather than raw data which reduces the great amount of time. The above experiment also clears that cluster-based ensemble handles categorical data very efficiently. There are different algorithms for clustering the categorical data. Each of one has its own advantages and limitations. The single popular clustering algorithm suffers from different problems like locally optimal clusters, dependence on initialization parameters, initial conditions, etc. The experimental result shows that

Table 5 Internal indies for cluster analysis of US census dataset

Index/algorithms	HC	PAM	DIANA	KM
calinski_harabasz	1895.251	2189.063	2192.295	2182.081
pbm	22,845,212	27,858,087	28,044,614	27,954,388
c_index	0.03754196	0.03003961	0.02995954	0.03023052
davies_bouldin	0.5110590	0.5092293	0.5093044	0.5096104
g_plus	0.01994852	0.01576455	0.01572195	0.01586500
silhouette	0.5401641	0.5598643	0.5604838	0.5601008
Compactness	649.3176	618.7823	618.5106	619.4695
Connectivity	15.98413	24.87937	24.74881	20.92302

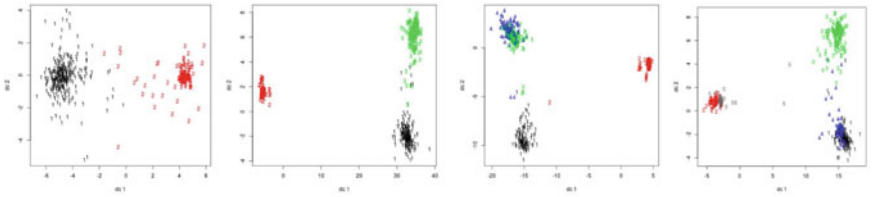


Fig. 1 Clustering results for US census data

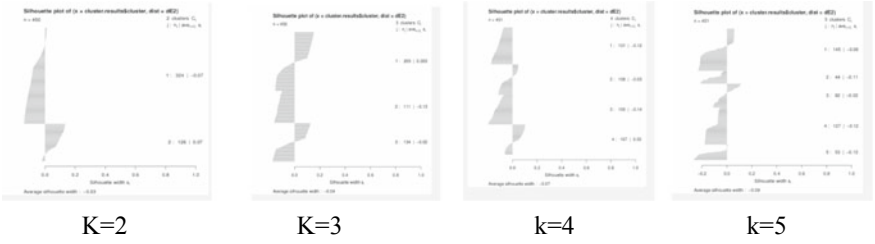


Fig. 2 Silhouette analysis for US census data

ensemble approach overcomes the limitations of individual clustering algorithms like K-mode, K-mean-ohe, hierarchical clustering (HC), PAM, DIANA, k-means and optimize the clustering result. The algorithms which perform well deserve to be the member of cluster ensemble. The relative performance of each algorithm is compared based on internal and external indices and the algorithms below a certain threshold are removed from ensemble. Thus, the strong members of the ensemble clustering take part in further process and improve the quality of clustering. Our future work will take advantage of these strong performing clustering members along with the proper classification algorithms [14] for optimizing the results in semi-supervised environment.

References

1. Pole, G., & Gera, P. (2016, September). A recent study of emerging tools and technologies boosting big data analytics. *Advances in Intelligent Systems and Computing (AISC)* (Vol. 413). Berlin: Springer.
2. Karthikeyani Visalakshi, N., & Arunprabha, K. (2015). Ensemble based distributed K-modes clustering. *International Journal of Engineering Research and Development*, *11*(3), 2278–2267.
3. Yang, Y., et al. (2016, May) Hybrid sampling-based clustering ensemble with global and local constitutions. *IEEE Transactions on Neural Networks and Learning Systems, 27*.
4. Strehl, A., & Ghosh, J. (2003, March) Cluster ensembles—A knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research, 3*, 583–617.
5. Yang, Y. (2011). Unsupervised ensemble learning and its application to temporal data clustering. In Ph.D. Dissertation, School *of* Computer Science, Manchester University, Manchester, U.K.
6. Fred, A. (2001). Finding consistent clusters in data partitions. In *Proceedings of 2nd International Workshop Multiple Classifier System* (pp. 309–318), Cambridge, U.K.
7. Fred, A. L. N., & Jain, A. K. (2002). Data clustering using evidence accumulation. In *Proceedings of 16th International Conference on Pattern Recognition* (pp. 276–280), Quebec City, QC, Canada.
8. Yang, Y., & Chen, K. (2010, July). Unsupervised learning via iteratively constructed clustering ensemble. In Proceedings of *International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8), Barcelona, Spain.
9. Topchy, A., Jain, A. K., & Punch, W. (2003, November). Combining multiple weak clusterings. In *Proceedings of 3rd IEEE International Conference on Data Mining, Melbourne* (pp. 331–338), FL, USA.
10. Fern, X. Z., & Brodley, C. E. (2004). Solving cluster ensemble problems by bipartite graph partitioning. In *Proceedings of the Twenty-First International Conference On Machine Learning* (pp. 36–42), Banff, AB, Canada.
11. Yang, Y., et al. (2011). Temporal data clustering via weighted clustering ensemble with different representations. *IEEE Transactions on Knowledge and Data Engineering, 23*(2), 307–320.
12. Qian, Y., et al. (2016, October). Space structure and clustering of categorical data. *IEEE Transactions on Neural Networks and Learning Systems, 27*(10).
13. Carreira-Perpiñán, M. Á., & Wang, W. (2013). *The K-modes algorithm for clustering*. ArXiv, abs/1304.6478.
14. Bommadevara, H. S. A., Sowmya, Y., & Pradeepini, G. (2019). Heart disease prediction using machine learning algorithms. *International Journal of Innovative Technology and Exploring Engineering, 8*(5), 270–272.
15. Reddy, M. V. J., & Kavitha, B. (2012). Clustering the mixed numerical and categorical dataset using similarity weight and filter method. *International Journal of Database Theory and Application, 5*(1).
16. Wang, F., et al. (2016). Empirical comparative analysis of 1-of-K-coding and K-prototypes in categorical clustering. In *CEUR Workshop Proceedings.*
17. Zengyou, H., Xiofei, X., & Shegchum, D, (2002, September). Squeezer: An efficient algorithm for clustering categorical data. *Journal of Computer Science and Technology*.
18. He, Z., Xu, X., & Deng, S. Improving categorical data clustering algorithm by weighting uncommon attribute value matches.
19. Xiang, Z., & Islam, M. Z. (2014). Hartigan's Method for K-Modes Clustering and its Advantages. In *Proceedings of the Twelvth Australasian Data Mining Conference* (AusDM 14) (pp. 25–30). Australian Computer Society Inc.
20. Meek, C., Thiesson, B., & Heckerman, D. (1990). UCI repository of machine learning databases.

21. Anila, M., & Pradeepini, G. (2017). Study of prediction algorithms for selecting appropriate classifier in machine learning. *Journal of Advanced Research in Dynamical and Control Systems, 9*(Special Issue 18), 257–268.
22. Schlimmer, J. (1981). UCI repository of machine learning databases.
23. Pole, G. S., & Potey, M.A. (2012). Highly efficient distributed indexing system based on large cluster of commodity machines. In *IEEE Conference on Wireless & Optical Communications Networks.*

New Avenues of Revenue in Indian Cinema

Sunitha Ratnakaram, Venkamaraju Chakravaram, Neelakantam Tatikonda, and G. Vidyasagar Rao

Abstract India's media and entertainment industry is of $23.9 billion in 2018 and is projected to reach $33 billion by 2021 [7]. Of this, $2.47 billion is the share of Indian film industry [7]. Film industry produces as many as 1600 films per year. 74% of revenue comes from home box office and 7% from overseas. Other sources of income account for as high as 19% ("Technology, Media and Telecommunications", 2016). In film industry, primary source of revenue till date is home box office collections. However, state of affairs is changing of late; where the film producers started searching for innovative avenues of revenue generation as the film making cost is rising hugely and the risk involved in movie making is going up. To list a few innovative sources of revenue goes like this; film-based merchandising, co-branding, brand associations, in-film advertising, home video, global marketing opportunities, digital platforms, broadband movie release, etc. Here comes the question, do Indian cinema really need some new sources of revenue when the industry is already making strong business? Of course, yes. Dependence on box office revenue alone means that successful outcome of the movie alone can save the producer. 2017 is one such year for "Bollywood" (name used for Hindi language movie industry), where as good as 90% of the movies are failures at box office ("Only Two Blockbusters", 2018). In 2018, a movie titled "2.0" costed film makers $82 million, and it went on to collect $115 million from the box office. Since 2017, there are the minimum seven to eight films that costed producers $30–50 million. Given rising production budgets, people involved in film production to distribution are safe only when the film fares well

S. Ratnakaram (✉) · V. Chakravaram
Jindal Global Business School, O. P. Jindal Global University, Sonipat, Haryana 131001, India
e-mail: sratnakaram@jgu.edu.in

V. Chakravaram
e-mail: vchakravaram@jgu.edu.in

N. Tatikonda
Wolkite University, Wolkite, Ethiopia
e-mail: neelakantamtati@gmail.com

G. Vidyasagar Rao
UCCBM, Osmania University, Hyderabad, India
e-mail: gvsraog@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_68

at the box office. However, on the downside if the film flops, entire money is lost. Having discussed the need of new sources of revenue for Indian cinema, this paper would give a detailed account on various other sources available in Indian cinema.

Keywords Indian cinema · In-film advertising · Film-based merchandising · Digital platforms · New sources of revenue · Co-branding

1 Introduction

Numbers furnished above and the business that is worth $2.47 billion surely bring us to a pertinent question, does Indian cinema needs any further new sources to generate additional revenue? Yes, of course. Let us understand the reason behind this answer. Primarily, Indian cinema depends on its box office collections as the primary source of revenue. And this dependence is as high as ninety percent. When these movies do well in terms of success at the box office, entire stake holders involved in the process gets their cake from the piece, and no one is at a disadvantage. However, when the movie becomes a dud at box office, all these stake holders or few of them might get affected adversely. If not all parties, few parties involved might become insolvent to if a big budget movie goes wrong. Let us see one such case.

Saaho (2019) is a regional film made in Telugu language and dubbed into Hindi and other regional languages. This movie is made on a budget of ₹350 crores (Wikipedia, 2019). After *Bahubali—the beginning* (2015) and *Bahubali—the conclusion* (2017) from the Bahubali franchise which made ₹2400 crores from the box office, Prabhas Uppalapati has become a bankable star, hence the budget. However, *Saaho* movie collected estimated ₹433 crores only from the box office making it a loser. With devastating performance in regional languages, it would have become an utter flop movie if not for the other avenues of revenue. *Saaho* made a prerelease business of ₹150 crores plus from selling audio, satellite, and online streaming platforms rights (Digital Native, 2018). This is not a standalone case, and there are many high profile disasters like *Bombay Velvet* (2015); Kites (2010); *Mohenjo Daro* (2016) to name a few; however, *Saaho* is surely one case where ill effects of box office got little manageable with other sources of revenue.

Most of the production houses in the western countries are evolved in their business processes, and they have learned to minimize dependence on box office for all their revenue proceeds. Alternative sources they bank upon are home video which provides constant stream of revenue and film-based merchandizing which opens up a complete new stream of opportunities for creating great markets for the stuff created based on the movies. This strategy brings down the risk played by the producer, some of these initiatives keep the producer afloat by protecting the cost of production even in case the movie not faring well at the box office. We suggest such similar initiatives could save Indian film industry where success rates of the films are plummeting year on year. Having alternatives streams of revenue can protect the sky rocketing production costs of the Indian movies.

Now we understand the backdrop that let us discuss about few new sources of revenue Indian cinema can try upon. There are many avenues like in-film advertising, co-branding, digital platforms, film-based merchandising, global marketing opportunities, brand associations, home video, broadband movie release, online streaming, to name a few.

2 New Sources of Revenue

2.1 In-Film Advertisement

This revenue source is not something very new in Indian cinema, and however, not something utilized to the fullest extent in its earlier version. This source is famously called as product placement in cinema too [6]. Mathur and Goswami [6] categorized product placement into two categories: one active placements and the other passive placement. Active placement is actor in the movie actively endorsing the product as part of portrayal of the role, and in passive placement, brand signage is placed in background.

Rishi Kapoor's debut movie *Bobby* (1973) is unique for many reasons apart from the main cast, one of those unique proposition is *Rajdoot*. In this movie, Rishi Kapoor visibly rides a Rajdoot bike of Enfield motors. Enfield motors used this cinema as its product launch platform benefiting both the parties. However, this is not the first instance of such kind. One noticeable instance from *An Evening in Paris* (1967) takes credit as the first. In this movie, lead actress sips her coke from a bottle of coke where she deliberately tries to unhide logo. Emami found its place in films as early as 1970s [6]. *Bobby* initiated a continuous on screen romance between brands and actors. From then on, we could witness many brand placement occasions in movies. Be it, Aamir Khan riding his BSA SLR bike in *Jo Jeeta Wohi Sikandar*, placing Mafatlal's advertisement in *Maine pyaar kiya* as part of a song sequence, and Stroh's beer explicit verbal and visual mention in *Dilwale Dulhaniya Le Jayenge* are a few to mention.

Compared to those earlier days of product placements, of late instances of in-film advertising certainly has increased. Producers too realized that creating in-film advertising opportunities for the brands can bring in additional revenue stream to the dwindling box office income. According to B. R. Chopra, "This kind of advertising helps producer to recover some of the costs of the film. In any case, in our films, we need to show characters driving cars, using mobiles, drinking tea, and using a number of other products, so why no benefit out of it?". This precisely makes the idea more realistic. Let us see an instance, Subhash Ghai the film maker of *Taal* chose to provide a product placement opportunity to coke in a prominent sequence and that one placement helped him recover 20% of the movie production cost. B. R. Chopra of course used this strategy. In his film *Baghban* Starring Amitabh Bachchan and Hema malini, he placed a hoard of brands from retail banker ICICI Bank, auto brand Ford

ikon, to Tata Tea. All these placements garnered him 3–4 crores before even getting on to box office battle [6]. We can mention many other notable instances like display of Hyundai Santro and Castrol Engine oil in *Chalthe Chalthe*, placement of 400 carats of real diamonds in Kaizad Gustad's *Boom*, Amirkhan receiving a phone call with Airtel ring tone in *3 idiots*, Hritik Roshan insisting to have Bournvita in *Krish*; CEAT tire placement in *Piku*, to name a few [10]. In-film advertising has become popular in the corporate world also due to its inherent advantages like availability of stars for a nominal cost comparatively, focused target audience, no zipping, zapping from the channels, etc.

2.2 Brand Association

In brand associations, movie is associated with a brand in all its promotional activities in various media. For example, Suraj Barjatya's *Hum Saath Saath Hai* associated its name with coke for ₹ 1.50 cores. *Mujhse Shadi karoge* with Britannia 50-50, *Kyon Ho Gaya Na* with Close Up, *Humtum* with Tata Indicom, *Lakshya* with Airtel, *Plan* with 8 PM whiskey are few more examples to mention here. In the recent times, brand associations have reached a different level with huge number of brand associations. This can be witness from *Veere di Wedding* movie's association with Air India, Amul, HSBC, Bharat Matrimony, Uber, Symbiosis, Lux, Apple, Tata Tigor, Tumi, Mydala, and Videocon [9]. *Sanju* made its association with Redmi Note 5 and Saregama Caravan. A movie made on a boxer's life *Mary Kom* attracted 20 brand associations [5].

2.3 Film-Based Merchandising

Film-based merchandising (FBM) can be defined as selling articles that are related to the movie's characters, incidents, and locations. These articles include toys, games, greeting cards, accessories, apparels, masks, videogames, etc. Videogames are very popular in today's film-based merchandising. FBM business is run through license mode of operation where producer of the movie gives the rights to licensee, i.e., manufacturer with royalty agreements. Here, licensee bears all the manufacturing and distribution expenses, whereas marketing activities will be conducted by both the parties combined.

History of FBM in India is of recent past, which started with movies like *Maine Pyaar Kiya, Roop Ki Rani Choron Ka Raja* in early 90s. FBM's success is majorly driven by the fate of the movie than the quality of the product which is evident from movies like *Badhaai ho Badhaai and Krrish*. Production house of *Badhaai ho Badhaai* released few greeting cards, stationery, and pens which did not fare well due to the failure of the movie in the market, and at the same time production house

of *Krrish* was able to sell entire kit of the movie on a very large scale due to the huge success of the movie.

FBM is a very successful story in Hollywood, but it is not so in India. There are many reasons for the same. Firstly, it's grey markets in unorganized retailing which sell duplicate products at very cheaper price compared to originals; secondly, hero worship which comes as an obstacle in selling the character; thirdly, limited experience of licensee in handling FBM which results either in too early or in too late release of FBM which cannot capture market; and lastly, movie releases with limited awareness created in the market. FBM is very much successful in the case of animated movies compared to the regular genre movies the reason being universal likeability as there no cultural, regional, or language barriers and most importantly directed at the major segment of children. Bahubali franchise broke this merchandising jinx in Indian films [4]. Director Rajamouli's franchise Bahubali banked upon its epic characters *Bahubali* and *Bhallaladeva* to create film-based merchandising starting from animation games, games, visual reality, TV series to a book series [8].

2.4 Co-branding

Co-branding in Indian cinema is not a recent phenomenon but today's marketer realized its importance. Co-branding is a concept where two companies come together for a specific project/cause/venture. When it comes to co-branding in cinema, marketer would like to tie up his brand with the movie either as a visual or clip or placement in the movie or it is used the lead characters in the movie. Co-branding is used by companies to have more of the visibility and awareness/image for the brand. May it be Bournvita in *Krrish*, Pepsi in *Kuch Kuch Hotha Hai*, or Tag heuer in *Don* worn by the lead couple serve the same purpose for the marketer. As earlier said, co-branding is not very recent in Indian cinema, actors in lead roles used to use bikes, apparel, accessories in the movie. In *Body guard*, film offers screen time to various brands like Audi Q7, Tupperware, Blackberry alongside its lead actors [2].

Nowadays, the corporate is looking at co-branding as many of the actors of the Indian cinema are brand ambassadors too for several companies. When the lead actor in a movie is one company's brand ambassador that company gets a greater advantage if it co-brands. When a company goes for co-branding, it should keep few things in mind. Firstly, suitability of the movie genre to the brand, secondly, suitability of the actor chose as ambassador in that movie, thirdly, suitability of the situation in the movie. We can witness very innovative forms of co-branding also these days, for example, Rajni Kanth starrer *Chandra Mukhi* movie opens with an advertisement of Tata Indicom, even after the interval movie succeeds only after the promo.

2.5 Global Market

Initially Indian cinema was confined to India itself. But, with the globalization and increased penetration of non-resident Indian in various parts of the world, demand for Indian cinema is growing across the world. Recently, Yash Raj Films was able to gain ₹30 crores from its *Veer Zaara.* Not only is this, *Kabhi Khushi Khabhi Gham, Kabhi Alvida Na Kehna, and Kal Ho Na Ho*, are few more names to mention here, which topped the international charts too. Indian cinema is doing well in traditional markets like US, UK, Canada, West Asia, South Africa, Mauritius, Fiji, Malaysia, Indonesia, Singapore, Hong Kong, Australia, and New Zealand. Indian producers are trying to exploit the potential available in new markets like Spain, France, Brazil, China, and Japan too. When it comes to Japan market, Super star from south, Rajani Kanth, is already well established with his movies like *Muththu*, *Baba*, and *Chandra Mukhi*. Recent story of overseas market opportunity is more exciting than earlier days. *Dangal* made ₹1300 crores from overseas market, whereas its indian box office revenue stands somewhere close to ₹600 crores. *Secret superstar* is another interesting story which made around 800 crores from global market, whereas its Indian box office collection did not even reach ₹100 cores. These two films opened up Chinese market for Indian films [1].

2.6 Home Video

As told earlier Indian cinema used to depend on box office collections only for all its revenue in the past. This overdependence on box office is making cinema to pile up huge losses. Home video is one alternative which can revitalize ailing Indian cinema to some extent. It works like this: most of the viewers are entering cinema theatre only when it is highly successful, the reason being the cost of movie viewing experience in theater has gone up and changes. Not only this, the people who could afford the price also can find time hardly for watching movie in a theatre. Home video is a good option for both the kinds of audience as it can serve both the purposes; it saves cost in case of budget class audience and time in case of affordable class audience. India's home video industry in the form of OTT is expected 11,000 crore industry by 2023. With Amazon Prime and Netflix entering Indian market, this segment has become strong and competitive. There are many small players too including Hotstar from the Star Network. Though each of them are coming up with original content, they still depend on main stream cinema for more content.

2.7 Broad Band Movie Release

As India is witnessing good penetration level of Internet in various regions, film producers have noticed a new opportunity in terms of broad band movies. Here, movie is available online for the viewers; they can either own the movie by downloading it from the respective site by paying a fixed amount or can view the movie once on broad band by paying a nominal amount. The Rajshree productions of Suraj Barjatya is the first company to launch this kind of service to the online audience. In fact, *Vivah* from Rajshree productions is the first movie to be launched through broad band which proved to be instant success. Later the company made all its titles to be available on net for sale. Reliance JioFiber with its first day first show plan is all set to make this segment more exciting [3]. This service is expected to be rolled out in the middle of 2020.

2.8 Digital Platforms

Producers are reaching technology savvy generation with this innovative platform, where they are developing mobile games and online games based on the content of the movie which creates hype in the market before the movie could release. Digital platforms are very much helpful for action-packed movies than the regular genre movies. There are many online games and mobile games developed for the movies like *Krrish, Kaante & Race*. Digital platforms are popular in regional movies too. Producers of *Jalsa* Telugu movie developed a mobile game which can be downloaded at free of cost. The moment digital platforms started streaming HD content this platform has become open to abundant opportunities. There are tens of apps that stream movies for free and in pay-per-view modes. Majority of the telecommunication corporations too have entered this segment to provide quality content to their premium segment customers. They are using the streaming services as part of their CRM program to build customer loyalty.

2.9 Satellite Rights

A film producer gives legal permission to broadcast their movie to a television channel for a limited period of time or for lifetime. With the influx of satellite channels in India, this has become another decent source of income for the producer. Satellite rights for the tech movie *2.0* is bagged by Zee network for 110 crores. As mentioned earlier *Saaho* received 70 crores from the sale of satellite rights. Though Indian television is still dominant by daily soaps, it does depend hugely on main stream cinema.

3 Road Ahead

One of the key emerging trends is co-production where major production companies from the world along with Indian major corporate are joining hands with film producer in making movie. Formerly Sony Pictures roped in Sanjay Leela Bhansali for a movie, but the most notices co-production venture was that of UTV's alliance with America's Fox Searchlight for Mira Nair's *The Namesake*. According to Subhash Ghai "Corporatized film production activities ensure respectability in transactions and operation and hence make it easier for film makers to source funds from the organized sector." Today, almost each movie is funded by at least one corporate as co-producer. This is giving more flexibility in terms of capital investment and better planning.

4 Conclusion

From the entire discussion, we can conclude that there are a variety of alternatives available with film producers to generate proceeds from film making, and all these ways can assure him of recovering production cost before the movie could release in the market. As of today, in-film advertising and co-branding together are generating 10–25% of production budgets, and all the other avenues discussed are also capable of engendering the major portion of the spent in recent future. Global business opportunity is the chief source of revenue which will become a driving force for the other options. Home video and broad band are not to talk less about kind of choices. When we talk of digital platforms, the world is being led by the same today so is the future. Corporatization of film production is really going to benefit the producer of today with all its advantages.

References[1]

1. Bollywood's Top Worldwide Earners | Gross Business of Over 200 Crores. (2019, April 10).
2. Kumar, A. (2014). Case study 5: Brands: The new co-stars in Bollywood. In *Marketing cases from emerging markets* (pp. 41–44). Berlin, Heidelberg: Springer.
3. Kumar. C. (2019, August 12). Jio First Day, First Show: JioFiber to air movies on release day.
4. Laghate, G. (2017, March 31).
5. Malvaniya. U. (2014, September 03). Film on Mary Kom attracts 20 brand associations.
6. Mathur, M., & Goswami, S. (2012). A study of consumer behavior and product placement in Indian cinema. *Annals of Management Research, 2*(1), 29–38.
7. Media and Entertainment. (2019, July).
8. Singh, A. (2017 March 15). How Bahubali franchise became the biggest film brand ever. Retrieved from: https://www.exchange4media.com/marketingnews/how-bahubali-franchise-became-the-biggest-film-brand-ever-68805.html

[1]Detailed and minor references available on request.

9. Mohan, N. (2018 December 31). Films that aced the art of brand association in 2018. Retrieved from: https://www.exchange4media.com/marketing-news/filmsthat-aced-the-art-of-brand-association-in-2018-93726.html
10. Vasan, A. (2018 March 06). Product Placements in Bollywood movies. Retrieved from: http://www.careerizma.com/blog/product-placements-bollywood-movies/)

e-Governance: Implementation of e-District Application

A Case Study in the District of Kangra, Himachal Pradesh

K. Rajeshwar and Manu Mahajan

Abstract Democratic governance mechanisms are becoming more receptive to the potential of Information and Communication Technologies (ICT) to achieve good governance in its implementation during the last two decades in India. In rural India, with the inception of NeGP initiatives and Digital India, the e-services are maximising and served at the doorstep of the citizen. It is also not only improving responsiveness of public service delivery mechanisms but also augmenting citizens' participation in governance mechanism especially through Lok Mitra Kendra's/Citizen Service Centre or through web portals. The Government of Himachal Pradesh has launched the e-District application under NeGD initiative. The average scores which have been arrived by aggregating the dichotomous replies given by each sample respondent are close to their respective maxima of the Digital penetration and LMK utility which are 3.34 and 3.47 respectively out of 4 and whereas for Portal features an average computed score is 4.41 out of a maximum score of 5 units which indicates that the actual implementation status of e-District portal at Kangra District is high. Through this case study, it is intended to study and document the process of implementation and effectiveness of e-District Application in the hilly district Kangra of Himachal Pradesh which is the largest populated district in Himachal.

Keywords e-Governance · ICT

K. Rajeshwar (✉)
NIRD&PR, MoRD, GoI, Hyderabad, India
e-mail: rajeshwar.nird@gov.in

M. Mahajan
SIRD, HIPA, Shimla, Himachal Pradesh, India
e-mail: mahajan.manu7@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_69

1 Introduction

e-District Mission Mode Project (MMP) is one of the 31 MMPs under National e-Governance Plan (NeGP). e-District MMP aims at electronic delivery of identified citizen centric services at District, Sub-Division and village level. It also envisages automation of workflow, backend computerization, and data digitization across participating departments. It integrates various digital database for online verification and complete automation of process. Nationwide, e-District services launched in 553[1] out of 672 districts including the 40 districts in pilot phase whereas Himachal Pradesh has rolled out e-District services in all 12 districts. The e-District architecture envisages leveraging of the core NeGP components of State Wide Area Network (SWAN), State Data Centre (SDC), State Services Delivery Gateway (SSDG) and Common Services Centres (CSCs).

The State Data Centre, HIM SWAN and Citizen Service/Lok Mitra Centres are the three pillars of the project. Citizen can avail these service under e-District MMP through multiple channels viz. Lok Mitra Kendra (LMK) Centres or Sugam Centres or Online through web login and from Department Office.

Himachal Pradesh, India is situated in the Western Himalayas with a difficult terrain and people have to travel long to avail the services from the public office. On another side, the good literacy rate i.e. 82.80% and a very good telecom density of ~148% helps the Government to provide the services at their door step. To provide the services effectively and efficiently, the Public Service Guarantee Act, 2011 is in place and the Digital Literacy Mapping[2] of the Government employees are in progress to build the capacity. Furthermore, the Lok Mitra Kendra/Citizen Service Centre are also established at the Panchayat level in the Public Private Partnership mode. The Government of Himachal Pradesh has introduced e-District application in the aegis of Department of Electronics & Information Technology (DeitY), Government of India and was conceptualized in the FY 2013-14[3] to enable seamless service delivery to the citizen at the district-level. In the state of Himachal Pradesh, the district administration is the primary delivery channel for G2C services. With the concatenation of e in the Governance process, the quality of the citizen will increase automatically and helps the administration for data mining with the centralized database for effective planning.

The Department of Information Technology is the nodal agency for the implementation of e-District MMP. Currently, 11 departments have hosted their services online which is further integrated with other databases like eParivaar and Digilocker. To execute, re-engineering of the internal processes have been carried out and the services are to be delivered to common man using the web portal or through the Common Service Centers (CSCs)/Lok Mitra Kendra's. In the F.Y. 2018–19, the district Kangra has registered highest number of 1.44 Million transactions where LMK has delivered the services at ground level with 1.42 M transactions. Maximum

[1] https://digitalindia.gov.in/content/status-mmps accessed on 3.5.2020.

[2] https://youtu.be/96Bg4DSEMF0 accessed on 10.12.2019.

[3] https://meity.gov.in/content/e-district-status accessed on 10.12.2019.

applications pertaining to Department of Revenue were submitted by using web login and LMK/CSC collected maximum of bill payments.[4] These figures have stated that the people of Kangra, HP have started to adopt the use of e-District application and LMK is serving best to the people. This success story will help to understand the process of rolling out, awareness and challenges and enabling factors undergone to execute the application.

1.1 Achievements

The Government of Himachal Pradesh has kick-started e-District application with 11 Departments and the project has also won the Skoch Order-of-Merits Award & Skoch e-Governance Gold Award 2016.

2 Scope of the Study

The study focuses on two main objectives which are mentioned below:

- To study the implementation of e-District application in the district Kangra of Himachal Pradesh.
- To study the digital penetration of services, usage of e-District application at grass root level and how those services are benefiting the beneficiaries.

3 Data Collection

Present study is based on quantitative as well as qualitative data collected from primary and secondary sources. The Methodology for collecting data for this study, visited to Kangra District, Himachal Pradesh. Primary data is collected from the beneficiaries, LMKs and department officials by administering the questionnaire tool and Secondary sources such as internet and department records.

To achieve the objectives of the study, the data collection tools developed for four categories such as beneficiaries, LMK's, Department Officials and Department of IT officials. The primary data will help in analysing & highlighting issues related to digital literacy awareness & knowledge, ease and convenience to users.

[4]Data accessed on url: http://edistrict.hp.gov.in/pages/staticSite/home.xhtml.

4 Sampling

Since the objective of the case study is to document process, achievements, challenges and also immediate impact, Kangra District was selected due to 100% implementation of e-District Application. In district Kangra, four blocks were selected viz., Indora, Baijnath, Lambagaon and Kangra because of highest and lowest level of literacy rate and sex ratio. The list of beneficiaries are collected from the Department and LMK officials during last two months whereas the LMK details are available online. Through tracer exercise, the data has been collected from all the selected blocks, total of 115, four LMKs from each block. From all the selected blocks, department officials such as Project Officer, Block Development Officers, Panchayat Secretaries and State IT officer from the Information Technology Department are primary source of information.

4.1 Period of Data Collection

Primary and secondary data from the selected beneficiaries list has been collected from the department and accordingly the Block and Gram Panchayats have visited to collect the data from them during November & December from state level IT department officials. Field visit has been conducted.

5 Overall Analysis

Taking the direction of these indicators into consideration, an index/score has been computed which reflects the extent of implementation of e-District Portal. It may be added that the index values have been arrived at by aggregating the dichotomous replies given by each sample respondent to each of the 13 indicators. Based on the nature of the indicators, they have been merged into the groups (dimensions) namely Digital penetration, Utility of LMKs and Portal features.

Table 1 shows the composition in terms of indicators of each dimension.

- The average scores of the Digital penetration and LMK utility are 3.34 and 3.47 respectively, where the maximum score is 4 and whereas for Portal features the average score is 4.41 out of a maximum score of 5 units. This indicates that the average scores are close to their respective maxima which further shows that in the perception of the sample respondents, the actual implementation status of e-District portal at Kangra District is very high (based on the 13 indicators considered).
- The total score on the 13 indicators is split into three groups viz., low, medium and high where low group means a score less than or equal to 8.0, moderate 9–10 and high includes those cases where the score value is greater than or equal to

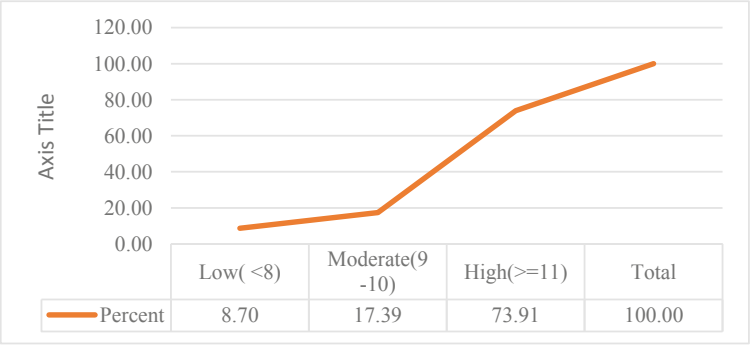
Table 1 Composition of the dimensions

S. No.	Index name	Indicators in the dimension
1	Digital penetration (4)	Usage of internet services, connected with social media, operation of digital device and holding of digital device
2	Utility of LMKs (4)	Consume less time at LMK, No rush at LMKs, LMKs adhere time frame wrt PSG Act 2011 and No extra charge for service provided
3	Portal features (5)	Getting of process updates of the portal, no issues in Payment Gateway, documents available in Digilocker portal, fulfilment of user requirement and simplification of form

11. Following this grouping, a graph is drawn which shows that there is a lot of clustering of scores in the high group accounting for 73.9%, followed by the moderate group with 17.4% and low group with hardly 8.7%.

Chart 3: Classification of scores into groups

- As each dimension is dependent upon varying number of indicators, to enable comparison across the dimensions, percentage scores are computed, Of the three dimensions, the Digital penetration gets the lowest at 83.5%, followed by LMK utility with 86.7%, and portal features 88.2%. The overall implementation score is 86.2 which shows a very high status of implementation of e-portal in Kangra district.
- The exercise serves two purposes—first, going by the total score the 13 indicators obtained, a clear endorsement of the fact that the indicators selected '*reflect the status of implementation*' of e-portal and second, the implementation status can be viewed and also assessed from the view point of three different dimensions which are Digital penetration, LMK utility and Portal features.



- The second part would enable the implementing agencies to take appropriate steps in respect of those dimensions which yield lower scores. From this point of view, this exercise is useful for adoption in like situations.

6 Salient Findings and Key Success Factors

The study involved a survey of four blocks in the district Kangra of Himachal Pradesh.

- As per the census records, the State is having a literacy rate of 82.8%, as per TRAI, the teledensity is 148% and as per e-PDS.HP 13.42% are BPL families in the district of Kangra of Himachal Pradesh. In the sample size of four blocks, 84% are graduates & above and 80%+ respondents from all four blocks are equipped with Digital Devices. 83.5% of the respondents are availing Internet services. This represents a very strong Digital penetration in the state.
- 81.7% of the respondents knows how to operate the digital device but 77.75% of them operate at basic/intermediate level only. It is also found that only 26% of them have undergone for any digital literacy programme and 69.6% exhibited their interest for hands-on training in Digital Literacy. In proportion to this the very less beneficiaries i.e. 8% availed e-Services through direct web login. This represents that the capacity building training programmes needs to be hosted at ground level which helps in the benefit of the rural people for availing e-Services.
- 100% of beneficiaries of sample size are aware of e-Services but 55.65% are aware that these e-Services are launched under e-District application. The friends & social gathering and Social Media are the major sources of penetration with 60.12% and 19.4% respectively. It also represents that more awareness programmes needs to be launched through Gram Sabha like public gatherings and using social Media extensively.
- The Citizen Service Centre, in the state it is known as Lok Mitra Kendra which are established at Panchayat level under the PPP mode. The CSCs are the channel between the e-Services and Citizens, in the sample size, 89% of the respondents availed these services through CSCs only. The beneficiaries are happy with the services of CSC as 95% found that they are delivering the services without consuming much time and 66.7% stated that they serve with in the time frames. It is also found that 80% of the beneficiaries have not found any display board and majority of LMK i.e. 53.3% stated that no monitoring is there. The Department of IT needs to ensure the sustainability and transparency of the CSCs. Moreover, the motivational factors are to be increased with more technical support and training to CSCs.
- The trust of the citizen on the portal is the major factor. In the analysis it is found that 73% of the respondents who have availed the services through web login trusted the data security of portal. The portal is continuously integrating with new databases and services in that order, the SSL certification and two-factor authentication needs to be incorporated in the application.
- The users found the application user-friendly, responsive and rated the application good. Moreover, 100% of the departmental officials rated the overall quality as good. The Department of IT is also planning to fuse the latest technologies to increase the outreach of the application in the rural masses. It resembles that in near future people will avail the services in a paperless and face less manner.

- For the progress of any district, the upliftment of Scheduled and Backward category is must. From the collected data it is evident that 78.3% i.e. the majority of the sample belongs to this category only which directly represents their upliftment of scheduled/backward category using e-Services.
- The sample size, also represented that majority of the beneficial belongs to marginal families i.e. 65% are in the income slab of Rs. 30K to Rs. 1 Lakh only meaning e-Services are actually serving the people who need it most and can't afford to visit Tehsil/Sub-Division/district offices for these services.
- The cyber Café available in the district is also charging minimal and the connection speed is also good helping the digital divide people to avail the e-services as per their needs.

7 Recommendations

The majority of the beneficiaries have reported their income in between 30K and 1L meaning they have very limited resources. Highest literate block Kangra is having 81.8% of APL cards and 27% of respondents are engaged in cultivation which is the highest no. of cultivators in comparison to other sample blocks. In the district, 87% respondents are equipped with digital devices out of which 91.8% are using smartphones where 83.5% of the respondents are using Internet services and 18.26% know how to operate digital devices at Intermediate or at advanced level. District is having a very good teledensity and internet subscribers. The district is having a very good awareness level i.e. 55.65% but due to the lack of digital awareness and capacity building they are availing the services through Lok Mitra Kendras. 89% of the respondents have availed e-Services through LMKs and only 8.38% have availed the services through direct/web-logins. In this scenario, awareness workshops and capacity building programmes on e-District portal, change & behavioural management trainings not only helps the people to get motivated and availed the e-services 24×7 which further helps in reduction of footfall in the Government setup for efficient and effective working. It is also suggested that the persons who can handle the devices at advanced level may act as master trainer at the ground zero.

The Lok Mitra Kendra are the foundation pillars established at the Panchayat level which act as an interface between the people and Government. From the sample it is found that the LMK officials are having a good understanding of the application, 80% are having a good computer knowledge, 66.7% found application less complex and 93.3% rated the good flexibility and responsiveness of the portal. Majority of the LMK is happy with the infrastructure but 86.7% face technical issues during providing the services that means a strong communication should be established between the LMK, DoIT and concerned line departmental officials which helps the LMK officials to serve the citizens effectively and efficiently. It is further suggested that the LMK should be monitored timely, proper capacity building activities are to be carried out with proper HIM SWAN connectivity.

8 Conclusion

e-District portal is the great success in the state of Himachal Pradesh as 84.6% departmental officials admitted that there is a reduction in the footfall in Government setup which reduces the pressure on Government machinery and completely satisfied with the e-District portal. The Lok Mitra Kendra's are the foundation pillars established at the Panchayat level which act as an interface between the people and Government and 89% respondents have availed the services through LMK only. This pushes the PPP mode and enables the beneficiaries to avail the services at their door steps. 66.7% of LMK officials found application less complex and 93.3% rated as flexible and responsive portal. 87% respondents are equipped with digital devices and 83.5% are internet subscribers in the sample size which means the digital penetration is good among society but 69.6% required the hands-on training on Digital Literacy. These training programmes may be executed with state SIRD/PRTI/ETC in cascading mode at panchayat level. This represents that the capacity of the respondents have to be built on Digital Literacy enabling them to avail the services as their own round the clock. The said model/project is the great success to reach the citizen to deliver the services. It is also strongly recommended from the above mentioned enabling factors that the said model may also be implemented in the North-Eastern states of India having almost the same hilly terrain as of Himachal helping the community to avail the services at their door steps round the clock.

9 Suggestions

It is evident from the analysis that the average scores are close to their respective maxima i.e. Digital penetration and LMK utility are 3.34 and 3.47 respectively out of the maximum score of 4 whereas for Portal features the average score is 4.41 out of a maximum score of 5 units which represents actual implementation status of e-District portal at Kangra District is high.

From the sample size, only 26% beneficiaries have attended any training programme on Digital Literacy and 69.6% showed their interest to undergo hands-on training. Moreover, awareness level has to be built among the rural masses as only 53.8% Departmental officials claimed that the citizen's are aware about these e-Services. It is also suggested that for gender mainstreaming, the role of the female needs to be encouraged at the village level.

The 93% of LMK officials are found the application flexible and 80% of them found it user-friendly but 86.7% have some technical issues while filling the application. And 66.7% got the technical maintenance over the portal which means a technical upgradation is required.

76.9% Departmental Officials are completely satisfied with the e-District portal and 84.6% of officials admitted that this portal helps them to reduce the public footfall in the Government setup and the citizens are now availing these services at

their doorsteps round the clock. It is therefore, strongly suggested that this portal and model may also be replicated in the North-Eastern states of India to deliver the citizen centric services at their door steps round the clock considering the difficult terrains of Himachal and North-Eastern states.

References

1. Alomari, M. K., Sandhu, K., & Woods, P. (2010). Measuring social factors in E-government adoption in the Hashemite Kingdom of Jordan. *International Journal of Digital Society (IJDS), 1*(2).
2. Dwivedi, S. K., & Bharti, A. K. (2010). E-governance in India—Problems and acceptability. *Journal of Theoretical and Applied Information Technology*.
3. Jehan, S. N., Nishantha, G. G. D., & Jehan, S. Q. (2010). E-governance initiative in Sri Lankan public service delivery. In *Institute of Electrical and Electronics Engineers (IEEE), 2010 ICACT*, Feb. 7–10.
4. Mofleh, S. I., & Wanous, M. (2008). *Understanding factors influencing citizens' adoption of e-government services in the developing world: Jordan as a case study*. http://www.dcc.ufla.br/infocomp/artigos/v7.2/art01.pdf.
5. Nabafu, R., & Maiga, G. (2012). A model of success factors for implementing local E-government in Uganda. *Electronic Journal of e-Government, 10*(1).
6. Sanyal, M., Das, S., & Bhadra, S. (2014). E-district portal for district administration in West Bengal, India: A survey to identify important factors towards citizen's satisfaction. *Procedia Economics and Finance, 11,* 510–521. https://doi.org/10.1016/S2212-5671(14)00217-2.
7. Singh, A. J., & Sharma, M. K. (2002). *District level e-governance initiative—A case study of Lok Mitra in Himachal Pradesh.*

Data Mining Techniques to Categorize Single Paragraph-Formed Self-narrated Stories

Md. Mahmudul Haque, Niloy Biswas, Niloy Saha Roy, Amzad Hossain Rafi, Sajid-ul Islam, Syeda Saiara Lubaba, Rakibul Islam, and Rashedur M. Rahman

Abstract In this age of natural language processing, most of the sentiment analysis tasks are done by polarization, for example, 0 for negative or 1 for positive of the given context/text. In some work, the tasks are done using fine-grained polarization, such as very negative or very positive. The proposed system of this paper includes the categorization of the paragraphs using its nature. All the paragraphs are self-narrated, and the number of words in those self-narrated paragraphs contains 50–4200 words. The paragraphs are categorized using three categorizations: "work stress," "bullying," and "sexual harassment" in both real and cyber worlds. Artificial neural network paragraph vectors, a distributed bag-of-words and distributed memory, are used to get the embedding of each paragraph and later for classification by data mining

Md. M. Haque (✉) · N. Biswas · N. S. Roy · A. H. Rafi · S. Islam · S. S. Lubaba · R. Islam · R. M. Rahman
Department of Electrical and Computer Engineering, North South University, Plot-15, Block-B, Bashundhara Residential Area, Dhaka, Bangladesh
e-mail: mahmudul.haque@northsouth.edu
URL: http://ece.northsouth.edu

N. Biswas
e-mail: niloy.biswas@northsouth.edu

N. S. Roy
e-mail: niloy.roy@northsouth.edu

A. H. Rafi
e-mail: amzad.rafi@northsouth.edu

S. Islam
e-mail: islam.saidul14@northsouth.edu

S. S. Lubaba
e-mail: syeda.lubaba@northsouth.edu

R. Islam
e-mail: rakibul.islam@northsouth.edu

R. M. Rahman
e-mail: rashedur.rahman@northsouth.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_70

techniques. The accuracy of each algorithm lies between 70 and 94%. The best model gives a 77.46% F1 score in the test set.

Keywords Data mining · Natural language processing · Classification · Document to vector · Supervised learning

1 Introduction

Automatic detection of sentiment from the story using deep learning tools and, in some cases, where the amount of dataset is low, machine learning tools are used and are the most current trends in recent years. Sentiment analysis itself is challenging due to the use of slang, ambiguity, sarcasm, and sometimes, these factors are misleading for both humans and machines as well. Analysis of sentiment is a series of methods and tools to extract information from language, such as views and opinions, as well as attitude [8]. This study is part of the ongoing research about categorizing the paragraph data based on sentiment analysis. The data are basically narrated stories described by individuals collected as a paragraph from different sources. All the stories are classified into three categories based on the story's sentiment: work stress, sexual harassment, and bullying. The scientific novelty of this paper is to categorize the given data from the described stories, which are self-narrated by the narrators. There are many ways to hurt someone, to bully someone, or to know one's mental state. Sometimes, a few words are not enough to express one's feelings. That is where this research shines the most. It is not necessarily bad words that hurt people all the time. If someone intends to hurt, bully, or put someone in stress, they may very well do that using nice words. What matters the most is the intention behind their words, to know the pattern of their behaviors and the consequences of them. The system that is built can detect such patterns dividing them in three above-mentioned categories. This research can be used to find the culprits behind these types of crimes. This research can be used in the complaint departments of any organization to sort out complaints of digital letters and how it should be treated. Also, this system can be used in online forums or social media to figure out such cases more effectively. The rest of the paper is organized into six sections. Section 2 presents the literature review. Section 3 contains a description of the dataset that is used in this research. A brief explanation of model architecture and methodology is in Sect. 4. Section 5 is for result analysis, and finally, the paper is concluded with the proposed future works in Sect. 6.

2 Related Works

In [19], the researchers utilized space-adjusted sentiment classification by taking notion from the objective area, and they formulated a novel in-depth adversarial mutual learning approach, including two groups of feature extractors, space discriminators, sentiment classifiers, and label probers. In [17], documented embedding models were used with cosine similarity, and a comparison between dot product and cosine similarity was presented. The authors also used the features with Naive Bayes weighted bag of *n*-grams and reported an accuracy of 97.42%. In [12], the Stanford Sentiment Treebank (SST)-5 was used to make a fine-grained classification using the BERT pre-trained model from Google. Inside BERT, the attention architecture processes the entire input in parallel to make the execution faster. After preprocessing, the filtered texts would go into BERT embedding; then, it would go through dropout regularization and then a SoftMax classifier layer to specify the output. The accuracy of this research was 94.7%. In paper [6], the authors expressed that the most regular use in non-flexible length features was bag-of-words though it accompanied two significant shortcomings. Another term was proposed by the authors called "paragraph vector," an unsupervised algorithm that gained from fixed-length writings. In "Learning Vector Representation of Words," the authors introduced the concept of distributed within vector-formed representation of words. Then, they trained the model to predict words randomly sampled from the paragraph in the output. The datasets they used were SST and IMDB datasets. They also did a test method on the information retrieval task. With SST paragraph to vector, the accuracy was 12.2% for binary and 52.3% for fine-grained classification. Moreover, with the IMDB dataset, the error rate was reduced to 7.42%. This paper [4] analyzed of damage to the flood of 2011 in Monterosso. The authors analyzed to trace the discursive strategies enacted to restore the tourist destination image from Web resources. They used a twofold approach: a linguistic and a sentiment analysis approach. They collected Web texts from 27 blogs. The authors performed double sentiment analysis from the contents of some selected Web sites and then considered the sentiments produced by the news reported on a disaster in social media. The authors reported that the predominant feeling became neutral over time but more emphasized negative feelings on the stages that immediately followed after the flood. The authors in [13] categorized positive, negative, and neutral sentiments, and tried to determine the opinions of users–parents toward an educational institution. They focused on two-point, alumni–parents' view after immediately joining the college, and then complaints related to colleges for those in the institute for a time, where two polar sides got the proper emphasis. They applied ANNIE application using the GATE tool for sentiment analysis to categorize positive, negative, and neutral sentiments. The researchers achieved this by extracting meaningful categories of information from the text documents. In [2], the researchers' core intention was to find out the correlation between denigration cyber-bullying and rumor—the proposed supervised model gained 83.4% accuracy using SVM with Gaussian radial basis function kernel. Their proposed model used a lexicon check to find out the presence of hateful words. In this paper [1], the

researchers attempted to detect cyber-bullying using Twitter data with psychological features such as personality, sentiment, and user emotions. The proposed classifiers were Naive Bayes, J48, and random forest.

3 Dataset

The dataset was collected from different sources which are: survey, open stories in blogs, social media public groups, and Kaggle's repository [5]. The survey dataset is collected by visiting the victims in person and also by online survey form. All the stories of the survey happened in Bangladesh. The open blog stories mostly originated in the USA and UK. The entire dataset was collected into a comma-separated value in (csv) file and later preprocessed using NLTK [9], Genism [16], and Genism's Doc2Vector [7] library using Python 3. The entire dataset has been classified according to three categories—0: work stress, 1: bullying, and 2: sexual harassment. The collected dataset contains information without revealing the contributors' personal identification. If any contributor provides their data using their name, then their names have been replaced using the character -X. In some cases, the locations which identify them or their relatives have also been replaced with the same character to avoid the model from creating some embeddings over those words. Another motivation behind replacing names is to prevent the model from creating any relation with the places; rather, the models are forced to create embeddings from the context of the stories as the dataset is limited. The separated numbers are kept the same; as in some stories, the numbers play a vital role. For illustration, in some "sexual harassment" data, the victim stated their ages or their academic classes to clarify that those incidents were also child abuse. The data distribution is self-narrated, which means their entire information is described by themselves or written by themselves. This also indicates that the dataset has come from a similar distribution for this paper's research purpose. The occurrences of data in the dataset are imbalanced which is evident from the count and Table 2. The number of total dataset is 273. The size of each paragraph is not equal. Some paragraphs contain less than 200 words, whereas some paragraphs contain words with more than 4000 words. The longest paragraph collected from an open blog story contains 4170 words, whereas the shortest paragraph contains less than 100 words (Table 1).

Table 1 Sample dataset format

Story	Category
My stress strikes when I find out that I am the less skilled guy in my…	0
At my old school kids would hit me and call me names…	1
I used to work for a call center and the men there were extremely…	2

Table 2 Total data count in each category with labels

Category	Total data count	Label
Work stress	79	0
Bullying (cyber and social)	98	1
Sexual harassment (cyber and social)	96	2

4 Methodology

The paragraph vector representation is proposed and described in [7]. The core intuition behind this is to let the machine learn from the context of the paragraph rather than the whole dataset. There are two types of paragraph vector representation used in this paper: "paragraph vector: distributed memory" (PV-DM) and "paragraph vector: distributed bag-of-words."

PV-DM: This method takes fixed-length samples from the context window. The paragraph vector uses a single document, whereas word vectors use all documents by either taking means or adding up both paragraph vectors. The next step is stochastic gradient descent, and gradients are obtained by backpropagation. This approach is similar to continuous bag-of-words from word2vec.

PV-DBOW: This is a different approach than that of PV-DM. In this method, it does not predict the next word, rather it applies a paragraph vector to classify entire words in a document. While training, it samples a list of words and forms a classifier to classify if the word belongs to the document so that word vectors can be learned. PV-DBOW resembles a skip-gram approach from word2vec.

4.1 Dataset Preprocessing

Tagged Document: Each document is made up of words, and those words are in Unicode string token format. The tags represent the class label of each document. The classes are in integer/ordinal encoded [14] where class values of each categories are: "work stress": "0", "bullying": "1", and "sexual harassment": "2".

Document to Vector: Genism's "Tagged Document" class replaces sentences as a list of words generated from the word2vec model of word to vector [10] implementation. "Tagged Document" receives tokenized texts, and the category of each document is known as class label. It uses the tagged data for the training and testing purpose of this research. After tokenizing, a single paragraph contains its tokenized words and its corresponding tag. In paragraph vector, distributed bag-of-words (PV-DBOW), Genism Doc2Vec class passes a parameter dm $= 0$, and for distributed memory (PV-DM), dm is 1. For PV-DBOW, the feature vector's dimensionality is kept 300×1, whereas the negative sampling of words is allowed at most 5 with minimum word count 2. The activation function used for this type of setup is hierarchical SoftMax [11]. The threshold for configuring most frequent words is randomly down-sampled to 0. Lastly, multiprocessing is used for better CPU utilization. Total

```
TaggedDocument(words=['male', 'customer', 'stalked', 'me', 'for',
'several', 'weeks', 'late', 'last', 'year', 'culminating', 'in', 'a
n', 'incident', 'where', 'he', 'followed', 'me', 'around', 'the',
'shop', 'shortly', 'before', 'close', 'casually', 'showing', 'me',
'what', 'appeared', 'to', 'be', 'gun', 'tucked', 'into', 'his', 'wa
istband', 'some', 'of', 'my', 'male', 'coworkers', 'told', 'me', 'w
as', 'eaggerating', 'or', 'that', 'the', 'guy', 'wasn', 'so', 'ba
d', 'another', 'my', 'manager', 'at', 'the', 'time', 'and', 'one',
'of', 'the', 'best', 'men', 've', 'ever', 'known', 'bought', 'me',
'knife', 'and', 'taught', 'me', 'how', 'to', 'use', 'it', 'when',
'told', 'him', 'what', 'had', 'happened'], tags=[2])
```

Fig. 1 A single paragraph with tokenized words and tag

epoch five is used during document vector training in both PV-DBOW and PV-DM. The initial learning rate for each model is 0.002. From there, the data and tag for training and validating are collected in vector form. K-fold cross-validation is used to find out the best split for the model whereas $K = 5$. In each training phase, training and validation split is: 218/55 (first threefold) and 219/54 (last twofold). In each fold, shuffling is used, setting the random state to 42 to make sure the model learns every class in its training phase. The fivefold cross-validation is used for the scarcity of available data in this domain. Figure 1 shows the tokenization with tags.

4.2 Embedding from Paragraphs

The vector dimension of each paragraph is 300×1 in PV-DBOW. The embeddings are in "tuple" [15] format. So, the data are immutable in this format. From the vector of PV-DM, the embeddings are represented similarly to the word2vec algorithm's continuous bag-of-word [10] which samples words from a given paragraph and then predicts the center word of the sample words. The vector dimension here is also 300×1. The paired model of PV-DM and PV-DBOW contains the document tags by keeping the inference of each type. Other temporary training data are deleted before concatenating two models. The vector dimension is 600×1.

4.3 Similarity and Dissimilarity of Paragraphs

The similarity measurement takes place via cosine similarity distance measurement. The reason behind choosing this similarity measurement for embeddings is to avoid bias due to the size of vectors as each paragraph's size is quite different from other paragraphs. To measure cosine similarity distance, a few paragraphs from the training phase are used. Table 3 contains the stories and labels, and Table 4 shows the cosine similarities among the data in Table 3.

Table 3 Sample training data with labels for measuring cosine similarity

Story	Category
I had debilitating migraines for three years… a week prior to leaving	0
The other part is that sense of worthlessness… and useless	0
My so-called friends in middle school used to call me Samara…	1
When I got admitted in a well-reputed college, I got bullied…	1
I'm a waitress, and I get harassed daily by customers trying to order…	2
Suddenly a male had come and was trying to touch…	2

Table 4 Distance measurement table

Data	Distance
Data 1 versus data 2	0.04
Data 3 versus data 4	0.09
Data 5 versus data 6	0.05
Data 1 versus data 3	0.30
Data 1 versus data 5	0.17
Data 4 versus data 6	0.31

In this paper, the cosine dissimilarity has been measured using SciPy [18] library, which measures the cosine similarity and then subtracts the value from 1. Among three embedding creations using the Doc2Vec model, PV-DM shows quite promising cosine similarity distance measurement. The best results are shown in Table 4. From Table 4, it is visible that the vectors of data with similar context have less distance among them, whereas the vectorized data with not similar context have a higher distance than a similar context.

4.4 Classification

For classification, 4 classifiers are used from 2 different genres. One genre is parametric learning, and the other is nonparametric learning. All the algorithms are implemented using Python's scientific learning API [3]. The algorithms are: logistic regression, Naive Bayes (Gaussian) classifier, decision tree classifier (entropy and Gini metrics), K-nearest neighbor classifier (Euclidean distance and Minkowski distance). As the dataset is small in size, these standard and lightweight data mining techniques are used in this research.

4.5 Parametric Learning Approach

Naive Bayes Algorithm using Gaussian Distribution: The prior probability of Gaussian distribution is not set, and thus the default variance smoothing is used from "sklearn" libraries.

Logistic Regression Algorithm: "Logistic regression" is used with sklearn's built-in linear model library. The number of CPU cores used is 1 with a more robust regularization of 1e5. Substantial regularization value is used to avoid any overfitting.

4.6 Nonparametric Learning Approach

Decision Tree Using Entropy: The maximum depth of the developed decision tree is set to 3. The minimum sample required to be a leaf node is set to the value of 5. The random state is 10.

Decision Tree Using Gini: All the other parameters are similar to the above statements of entropy criteria. The mapping of the decision tree with the paragraph vectors is based on the best split with the Gini value. In this section, the best decision tree with PV-DBOW is shown in Fig. 2. From Fig. 2, the "value" row in each node indicates the stored amount of observations in each node which fall into each of the three categories. The reason behind it is that if the features fall under the condition,

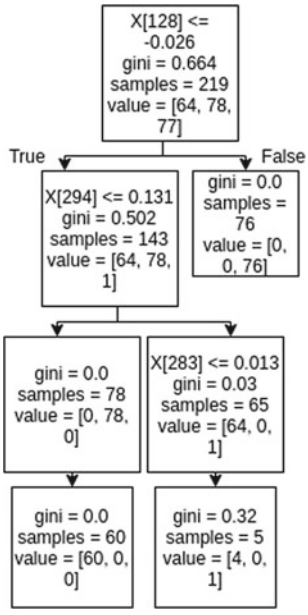


Fig. 2 Decision tree generated from PV-DBOW

then the edge value will be "true," otherwise the edge value will be "false." The value "samples" indicates the total value that is covered in that specific node.

KNN Using Minkowski and Euclidean Distance Metrics: The number of neighbors for KNN is set to 5. In the first algorithm, the distance metric is set to "Minkowski" distance metrics, and in the second one, "Euclidean" distance metric is used to determine the distance between two points.

5 Result Analysis

The training phase is performed more than once to achieve the best possible accuracy. Later in this section, there is a discussion about the test data and the best-performed model. The results which are described below are the reflection of the best models among the folds. For statistical analysis, F1, accuracy, precision, recall probability, and confusion matrix are used.

5.1 Parametric Learning Approach

Gaussian Naive Bayes: The highest accuracy of Gaussian Naive Bayes using both PV-DBOW and PV-DM on the testing dataset is about 88.8%, in the first fold. The confusion matrix is given in Fig. 3. The interesting fact is that Gaussian Naive Bayes with both PV-DBOW and PV-DM gives the exact similar confusion matrices, whereas the confusion matrix of the paired model shows more mis-classification than the other

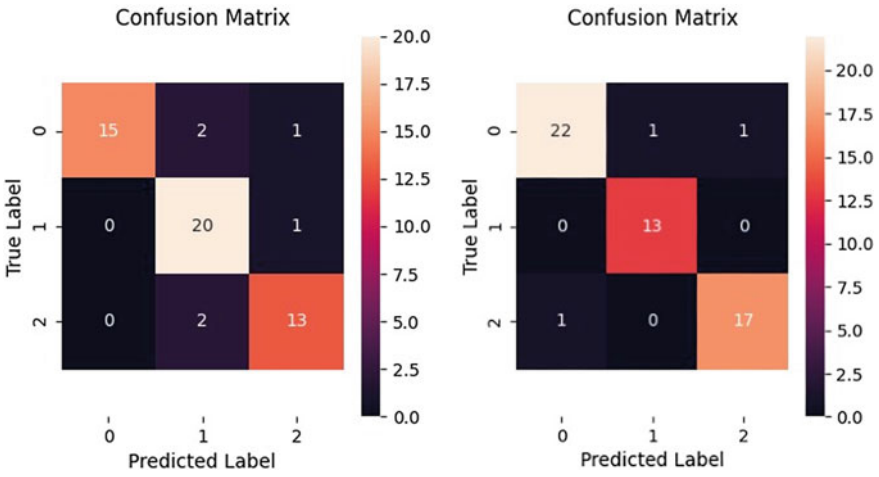


Fig. 3 Confusion matrix (heatmap): Gaussian Naive Bayes using PV-DM (at left) and KNN (Minkowski) using PV-DBOW (at right)

Testing Classification Report:

	precision	recall	f1-score	support
0	0.86	1.00	0.92	18
1	1.00	0.95	0.97	19
2	0.94	0.83	0.88	18
accuracy			0.93	55
macro avg	0.93	0.93	0.93	55
weighted avg	0.93	0.93	0.93	55

Testing Classification Report:

	precision	recall	f1-score	support
0	1.00	0.96	0.98	24
1	0.87	1.00	0.93	13
2	0.94	0.89	0.91	18
accuracy			0.95	55
macro avg	0.94	0.95	0.94	55
weighted avg	0.95	0.95	0.95	55

Fig. 4 Classification reports: logistic regression using PV-DBOW (at left) and decision tree using PV-DBOW and Gini indexing (at right)

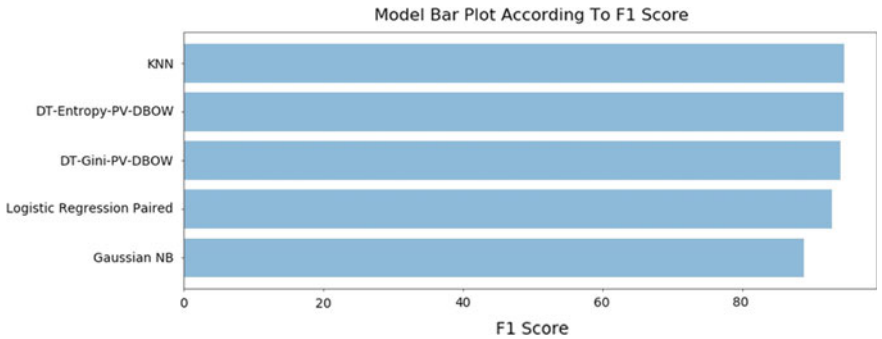


Fig. 5 F1 score bar plot of all the best models in validation phase

two. The paired model of PV-DBOW and PV-DM contains the trained model weights of both PV-DBOW and PV-DM. It has a consistent accuracy of 88.8%, and the F1 score is about 88.9 from fold numbers 2 to 5. The confusion matrix of all the folds is quite similar in terms of predicting false negative (FN), false positive (FP), true negative (TN), and true positive (TP). Fold number 4 has a quite promising score among all.

Logistic Regression: The F1 scores of PV-DBOW and PV-DM are 92.70 and 88.98%; the best score in this algorithm is 92.76%. From the confusion matrix, PV-DM with this algorithm performs poorly as it mis-classifies 6 data. The best one is PV-DBOW among three models. This model gives precision with 0.86 probability and recall probability is 1.00 in the class label "0", whereas precision 1.00 is in class label "1", recall is 0.95, and lastly, in class label "2", precision versus recall: 0.94 versus 0.83 which is shown in Fig. 4.

5.2 Nonparametric Learning Approach

Decision Tree: The decision tree using Gini index with PV-BOW gives an F1 score of 94%. The confusion matrix of this algorithm provides only three mis-classifications with the precision of 1.0, 0.87, and 0.94 in categories of "work stress," "bullying," and

Table 5 Best F1 scores in validation set with PV of different models

Model name	Paragraph vector type	Best F1 score (%)
KNN	PV-DBOW	94.54
Logistic regression	PV-Paired	92.8
Decision tree (entropy)	PV-DBOW	94.50
Decision tree (Gini)	PV-DBOW	94.00
Gaussian Naive Bayes	PV-paired	88.80

"sexual harassment" chronologically (Fig. 4). The decision tree with entropy metrics using PV-DBOW has an F1 score of 94.5% with similar mis-classification on the confusion matrix. However, in terms of precision and recall trade-off, it has given equal precision and recall value 0.94 on "sexual harassment" classification while giving less precision value (0.87) than recall value (1.00) in "bullying" categorization.

K-Nearest Neighbors: Both models with "Minkowski" and "Euclidean" distance have acquired 94.54% of F1 score in fold 3 and precision–recall ranging: work stress: 0.96 (precision), 0.92 (recall); bullying: 0.93 (precision), 1.00 (recall); and 0.94 (precision), 0.94 (recall) on sexual harassment using "Minkowski" distance metrics (Fig. 3). According to the F1 score on the validation set, the models' performances are given in Table 5 with the corresponding paragraph vector (PV).

5.3 Result Analysis on Test Set with Best Model

Judging by the overall confusion matrix output, F1 score, and precision–recall, decision tree with PV-DBOW with Gini has provided a promising performance on overall data. Nine data that are not part of the training phase are used to test this algorithm. All those 9 data are collected by conducting the victims in person; 7 out of 9 data have been classified correctly by this model and given in Fig. 6. F1 score is 77.46%. The data that it mis-classified are given in Fig. 7. The F1 scores of all the best models in the validation set are given in a bar plot form in Fig. 5. Though from F1 score, it looks like KNN has the upper hand, and both types of "decision trees" perform better in the confusion matrix. In terms of precision and recall trade-off, only decision tree with Gini using "PV-DBOW" has provided a promising score in unseen data. This model can only give a high score level of 77.46.

6 Conclusion and Future Work

This research is useful to gain attention on this aspect of text classification. The reason behind it is long and lengthy paragraphs/stories are also needed to be generalized or categorized to express the actual meaning of the entire story. To make the headfirst,

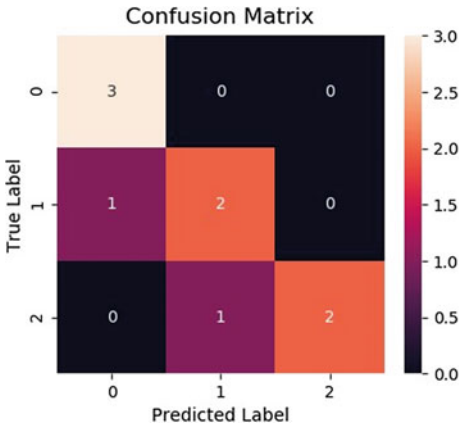


Fig. 6 Decision tree with PV-DBOW using Gini confusion matrix heatmap on test set

```
Story:
 i do 't let those who have abused me have control over me anymore i lost control of my life just about 7 years ag
o i went through things no little girl should at 13 14 15 16 17 everytime i 'd face one thing another terrible thi
ng would happen i wanted to stop fighting with everything i had but i did 't i kept sleeping on it and pushing thr
ough everything now im here engaged have my own home with my fiance and graduated with my associates degree in bak
ing and pastry i no longer fear to leave the house or go to work i have done everything i said i 'd never be able
to even though it still gets overwhelming and scary i did it my demons do 't own me anymore

Actual Label:  2

Model Predicted:  1

Story:
 i work in a ladies parlour when i was new to my office i had a toic relationship with one of my co workers she al
ways tried to find fault in my works and usually critisized me in front of other co workers this type of bullying
continued untill one day she got caught stealing customer 's money without their consent after she left everyone k
ind of got releaved as they also could not keep up with her

Actual Label:  1

Model Predicted:  0
```

Fig. 7 Decision tree with PV-DBOW using Gini mis-classified test data

this research will be the pioneer for next improved versions. The proposed algorithms have been trained and tested on a limited dataset. We plan to collect more data on three categories, and the dataset can be extended that will contain not only stories but also comments to find out the types as well as the offensive/insulting words from these three categories.

References

1. Balakrishnan, V., Khan, S., & Arabnia, H. R. (2020). Improving cyberbullying detection using twitter users' psychological features and machine learning. *Computers & Security, 90,* 101710.
2. Bhatia, M., & Sangwan, S. R. (2020). Debunking online reputation rumours using hybrid of lexicon-based and machine learning techniques. In *Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019)* (pp. 317–327). Springer.

3. Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., et al. (2013). API design for machine learning software: Experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning* (pp. 108–122).
4. Folgieri, R., Bait, M., & Carrion, J. P. M. (2016). A cognitive linguistic and sentiment analysis of blogs: Monterosso 2011 flooding. In *Tourism and culture in the age of innovation* (pp. 499–522). Cham, Switzerland: Springer.
5. Helpline of all sorts. https://www.kaggle.com/arnavbanerji/helpline-of-all-sorts.
6. Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. In *International Conference on Machine Learning* (pp. 1188–1196).
7. Le, Q. V., & Mikolov, T. (2014, May). Distributed representations of sentences and documents. https://arxiv.org/abs/1405.4053v2.
8. Liu, B. (2009). Sentiment analysis and subjectivity. *Handbook of natural language processing*. New York, NY, USA: Marcel Dekker, Inc.
9. Loper, E., & Bird, S. (2002). Nltk: The natural language toolkit. In *Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics. ETMTNLP '02* (Vol. 1, pp. 63–70). Stroudsburg, PA, USA: Association for Computational Linguistics. https://doi.org/10.3115/1118108.1118117.
10. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems. NIPS'13* (Vol. 2, pp. 3111–3119). USA: Curran Associates Inc. http://dl.acm.org/citation.cfm?id=2999792.2999959.
11. Mohammed, A. A., & Umaashankar, V. (2018). Effectiveness of hierarchical softmax in large scale classification tasks. CoRR abs/1812.05737. http://arxiv.org/abs/1812.05737.
12. Munikar, M., Shakya, S., & Shrestha, A. (2019). Fine-grained sentiment classification usingbert. arXiv preprint arXiv:1910.03474.
13. Patel, T., Undavia, J., & Patela, A. (2015). Sentiment analysis of parents feedback for educational institutes. *International Journal of Innovative and Emerging Research in Engineering, 2*(3), 75–78.
14. Potdar, K., Pardawala, T., & Pai, C. (2017). A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications, 175*, 7–9. https://doi.org/10.5120/ijca2017915495.
15. Ramalho, L. (2014). *Python tuples: Immutable but potentially changing—O'Reilly Radar*. http://radar.oreilly.com/2014/10/python-tuples-immutable-but-potentiallychanging.html.
16. Rehůřek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp. 45–50). Valletta, Malta: ELRA. http://is.muni.cz/publication/884893/en.
17. Thongtan, T., & Phienthrakul, T. (2019). Sentiment classification using document embeddings trained with cosine similarity. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop* (pp. 407–414). Florence, Italy: Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-2057.
18. Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2019). SciPy 1.0—Fundamental algorithms for scientific computing in Python. arXiv e-prints arXiv:1907.10121.
19. Xue, Q., Zhang, W., & Zha, H. (2020). Improving domain-adapted sentiment classification by deep adversarial mutual learning. In *AAAI* (pp. 9362–9369).

A Survey of Intelligent Agro-climate Decision Support Tool for Small-Scale Farmers: An Integration of Indigenous Knowledge, Mobile Phone Technology and Smart Sensors

Naledi Portia Thothela, Elisha Didam Markus, Muthoni Masinde, and Adnan M. Abu-Mahfouz

Abstract Food security in Africa and the rest of the globe has come under tremendous threat, meaning that agriculture which is the driving force behind many economies is under threat. Seventy per cent of the food produced in sub-Saharan Africa is produced in the rainfed small-holder agriculture, which in turn is the most devastated by any disasters experienced in the agricultural sector. Although the indigenous knowledge system has been used by many small-scale farmers as the basis for their day-to-day agricultural decision support, climate change and global warming have rendered this knowledge unreliable. Documented limitations on the isolated use of indigenous knowledge and modern scientific systems are the basis of this study. We investigate the effectiveness of the integration of indigenous knowledge interpreted through fuzzy inference systems, mobile phone and smart sensor technology with intelligence, on farmers decision support systems.

Keywords Fuzzy inference systems (FIS) · Indigenous knowledge · Decision support systems (DSS) · Cropping decisions · Small-scale farmers · Sensor technology

N. P. Thothela (✉) · M. Masinde
Department of Information Technology, Central University of Technology, Bloemfontein, South Africa
e-mail: portia.thothela@gmail.com

E. D. Markus
Department of Electrical, Electronic and Computer Engineering, Central University of Technology, Bloemfontein, South Africa

A. M. Abu-Mahfouz
Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_71

1 Introduction

Information and Communication Technologies (ICT) in agriculture may generally include technologies such as Global Positioning System (GPS), robotic systems, sensor technology and Geographic Information Systems (GIS). There is a plethora of these kinds of technologies already developed and put to a test in an effort to support agricultural practices and operations in order to improve production [1]. It should not come as a surprise therefore that so much research effort and resources have been put behind agricultural innovation. Agriculture is a mammoth economic driver in most countries and has the highest impact on the world's food security. Of all the food produced in sub-Saharan Africa, 80% is produced in the informal rainfed smallholder farms. Besides, out of 235 droughts reported worldwide between 2006 and 2015, 50% occurred in Africa. It is evident therefore that Africa is leading in terms of global droughts in comparison with the rest of the world [2]. What is more striking about these statistics is that these droughts have been reported to have been the cause of about 99% of deaths that occurred in Africa. Evidently, Africa is the worst affected by droughts, and small-scale farmers are the most tremendously affected by the most unpropitious disasters such as these, due to the absence of timeous warnings and relevant information [3].

It is important to differentiate small-scale farmers from the rest of the agricultural sector. According to [4], "A small-scale farmer is one whose scale of operation is too small to attract the provision of services he/she needs to be able to significantly increase his/her productivity". These farmers do not have access to equipment and machinery that can make their work easier. In fact, their work is labour intensive, and they rely on their human strength. They also have no access to any kind of ICT assistive tools due to the high costs related to such. Access to information such as weather forecast is another limitation for the small-scale farmer. Small-scale farmers mostly rely on their indigenous knowledge as a form of forecast and the basis for their daily decision support operation [5].

Why integrate indigenous knowledge into mobile technology, you may ask. Mobile phone technology is playing a colossal role in bridging the digital divide in Africa [6], i.e. the gap that exists between the demographic regions that have no access to modern ICT and Internet and those that don't. The mobile technology in comparison with other ICT technologies that is more affordable, user-friendly and less complicated in a sense that it has less infrastructure requirements. The mobile technology footprint in Africa stands at higher than 63% [7]. In Africa, mobile phone technology overtooks the use of landlines [6] to become the preferred mode of communication. In comparison with other available technologies in terms of performance and preference, mobile phone technology has become the technology of choice because of portability, high computing speed and power [8].

ICT in agriculture most often includes collaboration of two or more technologies. Agro-climate is a term used to refer to decision support tools in agriculture that incorporates climate information in crop management. Research has already shown that there is great potential in combining different ICT for agriculture to provide

smart technology in an effort to improve small-scale farmers' production [3]. In pursuit of such kind of smart technology, artificial intelligence (AI) and the creation of intelligent systems have become the new and most popular topic in technology.

This paper is aimed at developing a study that seeks to review and evaluate provision of the small-scale farmer with a portable, cost-effective, accessible and sustainable decision support tool. This would be done to enhance the small-scale farmers capability to make informed cropping decisions. The idea is to exploit technologies that have already been severally employed but never together. Incorporation of indigenous knowledge and to build this intelligent system will be evaluated to determine intensity of precision. Figures adapted from other papers shall be cited in the text and in the figures and tables section.

2 Mobile Phone Technology in Agriculture

2.1 Mobile Phone Technology Background

In 1947, when the transmitter was invented, an engineer at the Bell Labs sketched out the rough design for a standard cellular phone network, his name was Douglas H Ring. It would take 40 years for technology to catch up with that vision. Although Martin Cooper of Motorola is considered the inventor of the hand-held cellular telephone, he used technology developed by Bell Labs engineers. Since the very first invention, mobile phone technology evolved five generations and transformed the way we live. The first generation of mobile phone technology basically used analog radio signals. It was gradually replaced by the second generation which incorporated digital networks, and the first introduction of short message service (SMS) and multimedia messaging service (MMS). The third generation introduced multimedia application and Internet application with high connectivity speed. The fourth generation improved on the Internet connection speed, and the fifth generation is expected to solve most major smart phone issues and the security over the Internet.

2.2 Mobile Technology Is Sub-Saharan Africa

In terms of ICT indicators, though African has made a significant improvement in terms of Internet use, it is still far behind other counterparts from other regions [5]. In Africa, where the footprint of mobile phone stands higher than 63% [7], idiosyncrasies are towards mobile broadband as an Internet connectivity of choice because the cost of mobile broadband is fifty per cent lower than that of fixed broadband. According to [7], "Advancement in mobile phone technology has resulted in phones that can ably compete with personal computers of less than a decade ago; the devices (especially smart phones) are no longer used as mere phones for making/receiving

calls". The mobile phone added advantage is that it can act as both input and output device, owing to its highly technologically developed user interface.

The mobile phone technology has become a gigantic form of communication worldwide, and mobile phone networks have gained an explosive growth in the more recent past [6]. According to [6], mobile phone technology in sub-Saharan Africa has enhanced dissemination and retrieval of information for people, particularly in the rural communities. The author discovered that in Africa, mobile phones overtook the number of landlines. According to [9], "since the 2000s, mobile phone ownership has grown faster in Africa than anywhere else in the world". The author points to the fact that 83% of the population in Africa currently has mobile subscription. Mobile phones technology has also had an impact on the decrease of certain costs, resulting in improved functioning of certain sectors including agriculture, and there has been great improvement in the way business is conducted. Sub-Saharan Africa continues to lead the whole world in adoption of mobile money services [6, 10], and in sub-Saharan Africa, mobile industry continues to experience exponential growth reaching 367 million subscribers in 2015. Mobile phones are the most easily accessible and cost-effective technology available in South Africa. According to [11], "Africa has achieved a mobile phone penetration level much higher than that of computers". It is apparent therefore that it would be easy for any small-scale farmer operating anywhere to acquire one.

Mobile communication technology has become the most quotidian manner of transmitting data and services in a world that is constantly evolving. This drastic change gave rise to mobile applications [10]. The author suggests that mobile applications for agricultural and rural development possess potential for the advancement of rural development and can provide an affordable medium of information access to numerous people who were previously disadvantaged. Mobile phone technology seems to be bridging the gap in Africa between the rural demographic group with restricted access and the more urban demographic group with access to the latest ICT [6], owing to the fact that in comparison to other ICT mobile phone technology is much more affordable. Studies have also proved that if mobile phones were exploited properly they could improve the influence of sub-Saharan African farmers in the value chain [6].

According to [12], it is of the most utmost importance to explore intentions of small-scale farmers to adopt technological inventions that affect them. The study showed that expectancy of performance and effort, price value and trust proved critical to the adoption of such. In [6, 13], the author eluded to the fact that the farmers would have to be educated to realize the full potential of the mobile phone technology that is being offered in order to gain their buy in for the purpose of the adoption or acceptance of the ICT in their space. According to [6], studies have shown how technology and digitization impact on rural communities and that it can aid curb hinderances such as information access challenges for small-scale farmers.

2.3 Mobile Technology in Agriculture

According to [6], one of the most neglected global challenges that Africa faces is digital divide. However, the author eludes to the fact that this colossal challenge is being restrained by the adoption of mobile phones. In agriculture, technology-based decision support tools are typically software applications that commonly based on models that describe different processes in farming [10, 14]. "Mobile phone technology holds significant potential for advancing development" [10], but many technologies in agriculture, including those that are climate-smart are not being maximized because of low rate of adoption by small-scale farmers [1, 6]. Users for mobile technology in agriculture are also diverse and include several stakeholders. Mobile phone technology applications generally provide crucial information for rural small-scale farmers. DrumNet is one of the mobile applications used by Kenyan farmers, which has been instrumental in increasing their income by a third.

3 Integration of Indigenous Knowledge and ICTs in Small-Scale Farming in Africa

3.1 Indigenous Knowledge

The indigenous knowledge, which seems to be challenged currently by the ever-changing climate, is the local based (undocumented) knowledge that the small-scale farmers have employed for many years as a prediction tool and decision support management tool. It is somehow a legacy for the community and passed down from generation to generation [11]. Indigenous knowledge is normally based on the cultural ties to the natural environment, and the interpretation of the behaviour of nature around, including insects, animals and birds for example. This can include predictions such as rain or drought forecast, ploughing and planting season, harvest success and is based on environmental circumstances and events. These events can include, as an example, the yellowing of leaves on a crop, which will have a meaning attached to depending on the community. This has been the basis of the decision supporting tool used by small-scale farmers. The idea is not to confiscate the indigenous knowledge they are already familiar with, but to turn it into an enhanced decision supporting tool that is collaborated with modern science and technology.

3.2 ICTs in Small-Scale Agriculture

Information and Communication Technologies (ICT) in agriculture encompasses a plethora of technologies already in services [15], such as Global Positioning System (GPS), Geographic Information Systems (GIS), robotic systems and many more. In

South Africa, the agricultural industry is fractionated in two sectors. These include the large-scale commercial farmers and the small-scale rural farmers. The differentiation between the two was defined by the Department of Agriculture. According to [4], the Department of Agriculture states in their policy, "a small-scale farmer is one whose scale of operation is too small to attract the provision of services he/she needs to be able to significantly increase his/her productivity". The technology, mechanization and machinery available to support daily operations are in fact what separates the two, and the size has nothing to do with establishment of the difference. While state-of-the-art technology and sophisticated machinery is at the disposal of the commercial farmer for the support of daily running, the small-scale rural farmer will support his daily operations through intense labour, relying on their own strength and working with their hands. Hiring equipment is a luxury for the small-scale farmer.

3.3 Integration of Indigenous Knowledge in ICTs

Potential in collaboration of ICT to provide smart innovation for agriculture was uncovered in a study by [3]. The study proved that "the greatest impact of ICT can be expected in areas where farmers' actions are currently the primary limitation to their production". The author emphasizes that advancement in a single aspect of agricultural crop farming without taking into consideration all other relevant aspects such as soil tillage, geographic location, tolerance of environmental stress, soil type, rainfall, soil fertigation, pests and disease control and nutrition can be economically inefficient. The study concluded that it was critical to scale up on agricultural ICT services for small-scale farmers, but also that these services should be delivered at tailored local scale to for them to be relevant and to support the small-scale farmers' decision management.

A study reported in [16] highlighted the need for integration of indigenous knowledge in agricultural ICT development. According to the author, "farmers greatly value local experiential knowledge as they see it as having practical, personal and local relevance". It is therefore important to deliver information to small-scale farmers in a customized format that is relevant to them. In actual fact, studies have uncovered that more than 80% of farmers in countries such as Kenya, Zambia and Zimbabwe in the African continent are still dependent on their indigenous knowledge as a form of forecast on which their daily agricultural management decisions are based [2]. In the study by [17] the author quotes, "the effectiveness of the forecast information depends strongly on the systems that distribute the information, the channels of distribution, the recipients' models of understanding and judgement about the information sources, and the way in which the information is presented".

There is a strong relationship that exists between Indigenous knowledge and the decision support systems (DSS) that are based on scientific models. This relationship between the two areas of expertise in question has been accepted as rather complementary and not antagonistic. This manifested the basis of countless relevant and appropriate technologies that have been inaugurated to date [11]. Researchers are

also in one accord to the certitude that the integration between the two "can improve livelihoods", that means making a valuable difference in the lives of many. According to the author, "in order to build sustainable strategies, it is therefore important to take into account of, and learn from what the local people already know and do". As much as Indigenous knowledge and scientific models may harmonize, they are luminously contrasting expertise. Considering that Indigenous knowledge indicators have come under threat owing to the change in climate and global warming, it can be argued that science is needed to support the fading discipline.

Meteorological organizations in Africa have a colossal task with seasonal weather forecast generation. They are equipped with currently very few weather stations scattered over a vast area. This is an opportunity to introduce technologies that will help to bridge the gap and the indigenous knowledge of the local communities can contribute to the solution, giving them a sense of ownership of a system that is delivered. There is an opportunity that can be exploited in thorough unclouded comprehension of indigenous knowledge and appropriate integration with science, particularly in the distribution of weather forecasts to farmers in rural settlements.

In pursuit of the solution to the question, "does incorporating indigenous knowledge into the drought prediction tool improve the resilience and relevance to the countries in Sub-Saharan Africa", in a study by [17] a framework was developed for integration of mobile technology and indigenous knowledge for Africa. The idea is to enhance what the small-scale farmer also has, rather than replace it completely to curb resistance in adoption of the deliverable system. According to the study done by [6], the conclusion was that "perceived advantage and perceived usefulness influences mobile phone adoption negatively". This was attributed to the fact that most small-scale farmers only used their mobile phones for normal communication and not as an assistive tool for their agricultural activities. Incorporating what they already know, that their indigenous knowledge, could play a role in encouraging adoption of the deliverable system. This also encourages multi-purpose of their mobile phones.

4 Decision Support Systems for Small-Scale Farmers

"Agricultural decision support tools are typically software applications, commonly based on models describing various biophysical processes in farming systems and the response to varying management practices" [14]. Agricultural practices and operations are governed by decision management, which is in turn dependent upon or is supported by access to information such as climate forecast, soil fertigation, soil type and humidity. Crop farmers always need to be cognizant of all these factors to be able to make the right decision. According to [18], while either sparse or excess rain may kill crops, the adequate amount of rain spawns an ideal yield of crop. Hence small-scale farmers need timeous forecast on weather and climate. In sub-Saharan Africa, it is a gargantuan task for the rural small-scale farmers to just acquire seasonal climate forecasts even though it is constantly and regularly provided

by meteorological institutions, because of the manner in which the information is distributed [17].

A plethora of variable models and tools have been investigated to forecast various features of crop agriculture [15, 18–21]. It is lucid therefore, that predicting crops are complex enough to require a synthesis of various considerations to deduce the best agricultural practices that maximize productivity. The proposed system solicits a solution that merges and brings assorted aspects of crop agricultural decision support tools into convergence and disentangles small-scale farmer's decision support management.

Niche market would be a luxury for the small-scale farmer, and therefore because they do not specialize, they often plant an assortment of crops planted in the same field. Studies were conducted to determine how to deal with the challenge of water tolerance in intercropping [22] and to determine what crops would thrive if paired in intercropping process [23]. According to [24], enhanced productivity can become a consequence of intercropping that is balanced and well-managed. Successful intercropping happens when paired crops profit from one another in a positive exchange of nutrients or water source. Though small-scale farmers make use of intercropping, it is done from an illiterate and ill-informed place. The lack of information is the reason why it has not been exploited to full extent. According to [24], "farmers need an efficient, relevant and accurate way to evaluate data for specific management decision". This presents incorporation opportunity for a solution that can be incorporated into the proposed system.

According to [25], one of the challenges that are faced by small-scale farmers in developing countries is undiagnosed crop diseases or inefficiency in the diagnosis of crop diseases. This study also found that some of the drawbacks in existing systems in agriculture generally includes insufficient knowledge about the diseases and the farmers' lack of education. According to [14], a well-designed agricultural decision support system equips farmers with a facile and rapid way of comparing multiple scenarios for the management of crop production decisions. One example of such tools is CropARM, which assists farmers with management decision actions such as planting, by establishing risks taking into taking climate into consideration [14] have been designed.

5 Application of Fuzzy Inference Systems in Cropping Decision Support

5.1 Fuzzy Inference System

Since the introduction of the Internet of things (IoT), the need for intelligent systems has grown in popularity. Artificial intelligence is the ability that is built into a computer system to be able to apply the intricate function that a human brain uses to learn and make decisions. Machine-learning and fuzzy logic have become the more

widely used technological applications in the effort to produce intelligent systems. Fuzzy logic is a mathematical model in which logic depends on the degree of truth rather than a binary for where logic is either true or false. This kind of modelling is used in artificial intelligence, robotics, and business decision support systems. According to [25], fuzzy logic can be applied for decision-making for two reasons, "the rules are derived from expert knowledge that is described in natural language", and "it handles the vagueness and uncertainly inherent in the problem".

Fuzzy inference system can be used in this regard because it would take as input the information entered by the small-scale farmers as described according to their understanding. The information the farmer enters may be "fuzzy" or vague. Fuzzy logic is a multi-valued mathematical model that works on the degree of truth rather than either true or false. Fuzzy logic may be employed in image processing, decision-making and other artificial intelligence kinds of systems. Fuzzy logic magic may also be used over Android mobile. An intelligent framework was developed in the study by [25] for diagnosis of crop diseases. A fuzzy inference system (FIS), which is a rule-based system that makes use of fuzzy logic instead of normal Boolean logic, is exploited in this study. FIS provides an intelligent engine for synthesizing vague and uncertain information provided as part of the decision-making process.

The FIS process has three different stages as explained in [25], as shown in Fig. 1. First there is an input from the environment that must be fuzzified. This input passes through the inference engine where the fuzzy rules are applied using the fuzzy logic. A mapping according to the degree of truth is used to get a fuzzy value. This value is used in one of two methods of defuzzification, centroid or middle of maximum (mom). The following formulas apply:

$$\text{Centroid } z^* = \int \mu(x)x\mathrm{d}x/\mu(x)\mathrm{d}x$$

$$\text{Mom } z^* = a + b/2$$

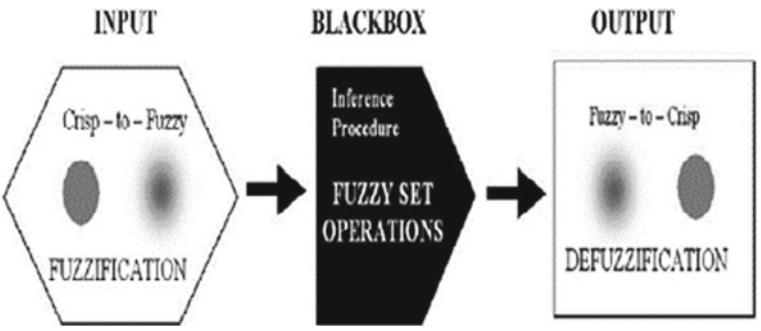


Fig. 1 Defuzzification through fuzzy inference engine

5.2 Cropping Decision Support for Farmers

According to [26], "agricultural statistics reveal interannual variations in the proportions of crops". This can be attributed to climate change and global warming. The variation in the climatic change becomes difficult to keep track of if it is not properly documented and causes environmental uncertainties. It is very important to understand these climatic variations since it affects the agricultural decision the farmers have to make regarding their crop that directly affects the quality and quantity of the harvest [26]. Agricultural farming is also impacted by economic circumstances, access to credit and markets [26]. With such a multitude of external factors affecting operations in farming, small-scale farmers find themselves having a mammoth task of having to make strategic, operational, tactical, long-term and short-term decisions [26]. Without relevant information, it becomes a daunting task for small-scale farmers to make informed decisions.

Agricultural farming is also impacted by economic circumstances, access to credit and markets [26]. With such a multitude of external factors affecting operations in farming, small-scale farmers find themselves having a mammoth task of having to make strategic, operational, tactical, long-term and short-term decisions [26]. Without relevant information, it becomes a daunting task for small-scale farmers to make informed decisions.

5.3 Applying the Fuzzy Inference System to Cropping Decisions

In the study by [25], a diagnostic expert system with artificial intelligence to diagnose crop diseases. The architecture for this system used fuzzy logic as a backend decision engine and was designed to work with Android mobile technology employing jFuzzylite library. Figure 2 depicts the architecture of the mobile application developed in the study [25]. According to the author, fuzzy logic was used for decision support in this frame for two main reasons. The one reason is that the rules are acquired from expert knowledge articulated in simple language and fuzzy logic does a powerful representation for linguistic knowledge. The other reason is that it can handle uncertainty of the input thereby providing the farmer with the opportunity to be as intuitive as possible in their input guesses. This also explains why fuzzy logic was employed in the study [17] in the development of the ITIKI framework shown in Fig. 2.

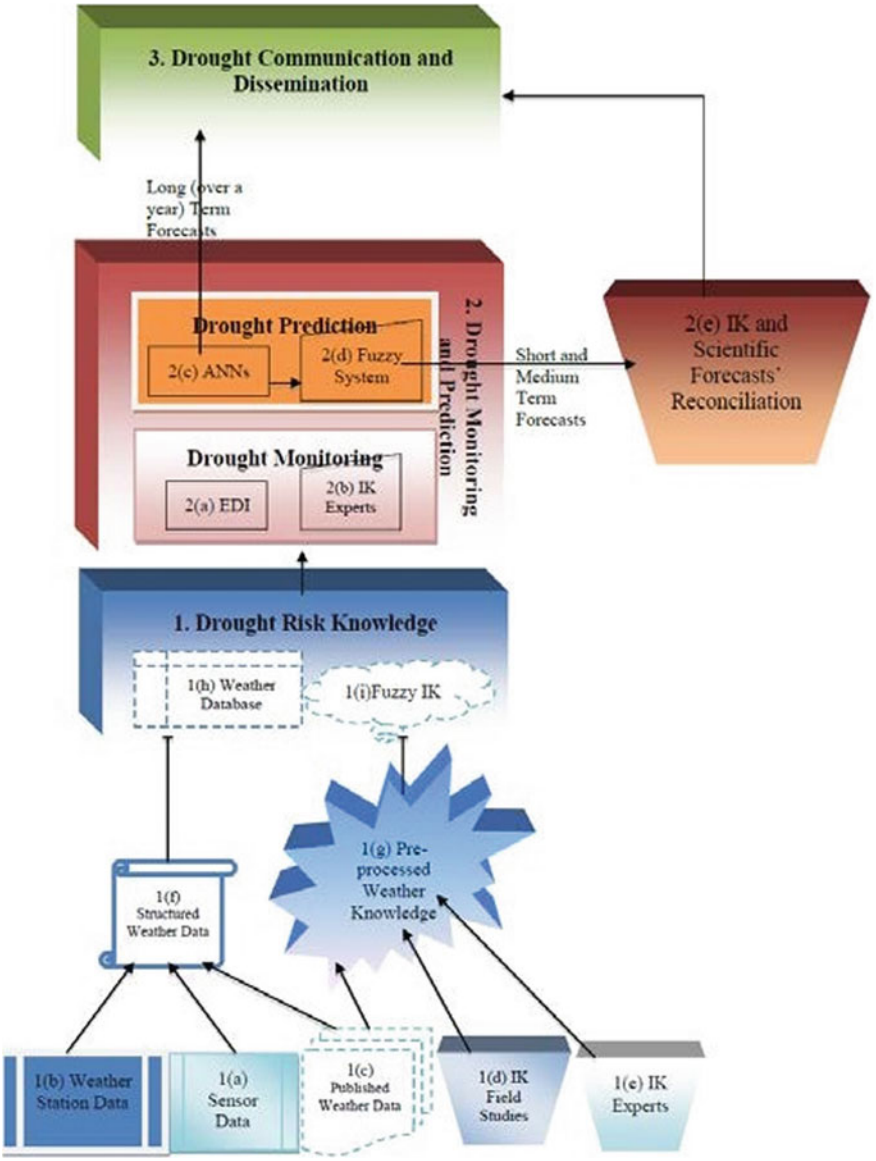


Fig. 2 ITIKI: Integrated Drought Early Warning System Framework

6 Comparison and Discussion

The ITIKI system did well in incorporating the indigenous knowledge, Artificial neural networks and the sensor technology to assist small-scale farmers with drought prediction [17]. This is a crucial solution considering that small-scale farmers operate

in rainfed agriculture. Rainfall is but only one part involved in agricultural decision-making. This prediction tool has been instrumental in the improvement of the farming decisions in sub-Saharan Africa. The challenge that the system has is that it does not provide the farmers with holistic agricultural information required to make strategic, tactical and operational management, such as decisions on fertilizer options, choice of crop for different conditions provided, agricultural drought, and the intercropping potential to exploit water resilient crops for maximum benefit.

The advantage of the ICT solution in the Intelligent mobile application system applied in the study by [27] is that it allows for uncertainty of input to go through a fuzzy inference engine to provide a prediction on crop diseases. The system also employs the use of mobile phone technology, which is a kind of technology that is more accessible to the small-scale farming technology in sub-Saharan Africa. The challenge the system has is that it does not allow for expert authentication of the output at runtime. The system does not cater for a more encompassing solution towards agricultural challenges that the rural small-scale farmers encounter, making it a single target problem-solving solution.

While SIMAGRI, also an agro-climate system was aimed at assisting strategic and tactical decision in the production of crops, it uses historical weather to predict climate crop yield by running a simulation of "what if" scenarios [28]. Advantage of the system is that it takes a lot more management disciplines into consideration, such as planting dates, fertilizer type and environmental conditions, but on top of that it takes user input. The disadvantage is that the system simulates yield predictions in a graphical manner which is not what small-scale farmers sub-Saharan Africa need.

CropARM is one other tool that was developed in the study by [14] in order to assist users to establish a framework of risk. It used the APSIM model to simulate scenarios using climate conditions and management actions. Though this system is available online and assists in management decisions that generated very close to predicted yields, it is not accessible to small-scale farmers in rural areas. In the study by [26], one of the limitations was determining the influential factors to cropping plans, meaning that there was no clear indication of the reasoning behind tactical decision. The author eluded to the fact that the small community based their choices on what the society in general was doing. This clearly shows that without enough resources and information, small-scale farmers agricultural decisions could clearly be hampered.

Table 1 highlights advantages and disadvantages of four agricultural solutions that already exist. All the solutions discussed had very successful results in their respective investigations, but this paper set one's sight on the emphasis of the target users of the envisaged system under scrutiny. The disadvantages are therefore highlighted in the comparison table. SIMAGRI is a highly specialized system that supports multiple disciplines of agricultural management using historic data, making it a daunting task to enable the use for decision management for small-scale rural farmers. Though CropARM establishes a framework of risk in agriculture by merging climate scenarios with agricultural management actions, it provides a complicated interface for output or results which would make it not user-friendly for the rural small-scale farmer. This makes both solutions highly inaccessible to small-scale farmers due to

Table 1 Technology comparison

ICT solution	Technology used	Advantages	Disadvantages
SIMAGRI	DSSAT	• Supports multiple strategic and tactical agricultural decisions • Uses tercile seasonal climate and climatology for risk analysis	• Can be exigent for non-experts to apply for the purpose of agricultural decision-making
CropARM	APSIM	• Help establish a framework of risk by incorporating climate scenarios and management actions • Regular updated climatic streams	• Inaccessible to small-scale rural farmers • Output method is not user-friendly for the small-scale farmers
Intelligent Mobile App	Fuzzy logic Mobile phone technologies	• Able to interpret vague inputs • Uses mobile phones technology and android OS • Incorporates local language for the for the farmers	• For diagnosis of crop diseases only • Does not take many other agricultural management components into account
ITIKI	Fuzzy logic Mobile phone technologies ANN	• Early warning prediction for droughts • Able to interpret vague inputs • Incorporates Indigenous knowledge and is user friendly for small-scale farmers • Built to be user-friendly for small-scale farmers	• Does not take many agricultural management components into account, it is only a drought prediction tool

the lack of education in rural areas. It is apparent therefore that these solutions are not suited for target user in this case.

The intelligent mobile application for diagnosis of crop diseases is only utilitarian in a case where crop health is in question, which is passable given the disaster that small-scale farmers experience when crop diseases go undiagnosed. This brings us to the question, 'what about cases where there is no crop health scare'. How will this tool be of any use to a small-scale farmer? The answer is then that it would be impractical. This disqualifies the solution for small-scale farmers because it is inaccessible to them. The ITIKI seems at this stage to be the only solution that is focused mainly on assisting small-scale farmers. It incorporates the small-scale farmer's indigenous knowledge in a drought prediction tool, allows for vague user input and was developed for small-scale farmers, making it a more user-friendly tool. Though this is only a drought prediction tool which is a bit of a disadvantage, it provides a foundation for

a decision support tool that is more encompassing of agricultural management and management actions. With the prescription of the foundation, therefore an intelligent agro-climate decision support tool can be instituted.

7 Conclusion

In the various research experiments that have been performed, it has been substantiated through research that certain limitations can be attributed to scientific models that certain decision support systems are based on [29], even though they can also be very crucial in agronomics. This furnishes opportunity in exploring innovation in technology that integrates scientific models, agro-climate, mobile phone technology, indigenous knowledge and artificial intelligence into a decision support system that supports an intelligent crop prediction tool. Research has also shown that the sub-Saharan African small-scale farmer is in desperate need of a sustainable, accessible and cost-effective, comprehensive decision support tool to support daily operation, for short term and long term, that takes into consideration the knowledge that they possess and can bring to the table [5, 7, 11, 30]. This decision support tool must also be acceptable by the small-scale farmer more than just imposed on them, that there may be a level of ownership they can take some pride in. The proposed solution seeks to amalgamate and merge divergent ICT to provide a solution more fitting for the sub-Saharan small-scale farmer using the ITIKI framework (ITIKI Plus) [30].

Acknowledgements This research was supported by the Council for Scientific and Industrial Research, Pretoria, South Africa, through the Smart Networks collaboration initiative and IoT-Factory Program (Funded by the Department of Science and Innovation (DSI), South Africa). The authors also would like to appreciate the support of the Department of Electrical, Electronics and Computer Engineering and the Department of Information Technology at the Central University of Technology.

References

1. Westermann, O., Förch, W., Thornton, P., Körner, J., Cramer, L., & Campbell, B. (2018). Scaling up agricultural interventions: Case studies of climate-smart agriculture. *Agricultural Systems, 165,* 283–293.
2. Akanbi, A. K., & Masinde, M. (2018). Towards the development of a rule-based drought early warning expert systems using indigenous knowledge. In *International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)*, Durban, 2018 (pp. 1–8).
3. Amarnath, G., Simons, G. W. H., Alahacoon, N., Smakhtin, V., Sharma, B., Gismalla, Y., et al. (2018). Using smart ICT to provide weather and water information to smallholders in Africa: The case of the Gash River Basin, Sudan. *Climate Risk Management, 22*, 52–66.
4. Kirsten, J. F., & Van Zyl, J. (1998). Defining small-scale farmers in the south african context. *Agrekon, 37*(4), 551–562.
5. Masinde, M. (2015). MAS-DEWS: A multi-agent system for predicting Africa's drought. In *International Joint Conference on Neural Networks*, September 2015.

6. Kabbiri, R., Dora, M., Kumar, V., Elepu, G., & Gellynck, X. (2018). Mobile phone adoption in agri-food sector: Are farmers in Sub-Saharan Africa connected? *Technological Forecasting and Social Change, 131*, 253–261.
7. Masinde, M. (2014). IoT applications that work for the African continent: Innovation or adoption? In *Proceedings—2014 12th IEEE International Conference on Industrial Informatics* (pp. 633–638).
8. Adepu, S., & Adler, R. F. (2016). A comparison of performance and preference on mobile devices vs. desktop computers. In *2016 IEEE 7th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2016*.
9. Jeffers, V. F., Humber, F., Nohasiarivelo, T., Botosoamananto, R., & Anderson, L. G. (2019). Trialling the use of smartphones as a tool to address gaps in small-scale fisheries catch data in southwest Madagascar. *Marine Policy, 99*, 267–274.
10. Qiang, C. Z., Kuek, S. C., Dymond, A., & Esselaar, S. (2011). *Mobile applications for agriculture and rural development*.
11. Masinde, M., Bagula, A., & Muthama, N. (2013). Implementation roadmap for downscaling drought forecasts in Mbeere using ITIKI. In *International Telecommunication Union—2013 Proceedings of ITU Kaleidoscope: Building Sustainable Communities, K 2013* (pp. 63–70).
12. Beza, E., Reidsma, P., Poortvliet, P. M., Belay, M. M., Bijen, B. S., & Kooistra, L. (2018). Exploring farmers' intentions to adopt mobile Short Message Service (SMS) for citizen science in agriculture. *Computers and Electronics in Agriculture, 151*, 295–310.
13. Meijer, S. S., Catacutan, D., Ajayi, O. C., Sileshi, G. W., & Nieuwenhuis, M. (2015). The role of knowledge, attitudes and perceptions in the uptake of agricultural and agroforestry innovations among smallholder farmers in sub-Saharan Africa. *International Journal of Agricultural Sustainability, 13*(1).
14. Phelan, D. C., et al. (2018). Advancing a farmer decision support tool for agronomic decisions on rainfed and irrigated wheat cropping in Tasmania. *Agricultural Systems, 167*(September), 113–124.
15. Manjula, A., & Narsimha, G. (2015). XCYPF: A flexible and extensible framework for agricultural Crop Yield Prediction. In *2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO) ISCO 2015*.
16. Šūmane, S., et al. (2018). Local and farmers' knowledge matters! How integrating informal and formal knowledge enhances sustainable and resilient agriculture. *Journal of Rural Studies, 59,* 232–241.
17. Masinde, M. (2015). An innovative drought early warning system for sub-saharan Africa: Integrating modern and indigenous approaches. *African Journal of Science, Technology, Innovation and Development, 7*(1), 8–25.
18. Ahamed, A. T. M. S., Mahmood, N. T., Hossain, N., Kabir, M. T., Das, K., & Rahman, F., et al. (2015). Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh. In *2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)—Proceedings*.
19. Gandhi, N., Petkar, O., & Armstrong, L. J. (2016). Rice crop yield prediction using Artificial Neural Networks. In *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)* (pp. 105–110).
20. Fan, W., Chong, C., Xiaoling, G., Hua, Y., & Juyun, W. (2015). Prediction of crop yield using big data. In *2015 8th International Symposium on Computational Intelligence and Design (ISCID)* (pp. 255–260).
21. Baruah, R. D., Roy, S., Bhagat, R. M., & Sethi, L. N. (2017). Use of data mining technique for prediction of tea yield in the face of climate change of Assam, India. In *Proceedings—2016 International Conference on Information Technology (ICIT) ICIT 2016* (pp. 0–4).
22. Chibarabada, T. P., Modi, A. T., & Mabhaudhi, T. (2017). Nutrient content and nutritional water productivity of selected grain legumes in response to production environment. *International Journal of Environmental Research and Public Health, 14*(11).

23. Chimonyo, V. G. P., Modi, A. T., & Mabhaudhi, T. (2016). Assessment of sorghum–cowpea intercrop system under water-limited conditions using a decision support tool. *Water SA, 42*(2), 316–327.
24. Chimonyo, V. G. P., Modi, A. T., & Mabhaudhi, T. (2016). Water use and productivity of a sorghum-cowpea-bottle gourd intercrop system. *Agricultural Water Management, 165,* 82–96.
25. Toseef, M., & Khan, M. J. (2018). An intelligent mobile application for diagnosis of crop diseases in Pakistan using fuzzy inference system. *Computers and Electronics in Agriculture, 153,* 1–11.
26. Jahel, C., Augusseau, X., & Lo Seen, D. (2018). Modelling cropping plan strategies: What decision margin for farmers in Burkina Faso? *Agricultural Systems, 167*, 17–33.
27. Niebel, T., et al. (2018). An intelligent mobile application for diagnosis of crop diseases in Pakistan using fuzzy inference system. *World Development, 167,* 52–66.
28. Han, E., Baethgen, W. E., Ines, A. V. M., & Mer, F. (2018). SIMAGRI: An agro-climate decision support tool. *Computers and Electronics in Agriculture, 161*, 241–251.
29. Vijayabaskar, P. S. (2017). *Crop prediction using predictive analytics* (pp. 370–373).
30. Masinde, M., & Thothela, P. N. (2019). ITIKI Plus: A mobile based application for integrating indigenous knowledge and scientific agro-climate decision support for Africa's small-scale farmers. In *2019 the 2nd International Conference on Information and Computer Technologies (ICICT)*, Kahului, HI, USA, 2019 (pp. 303–309).

Effectiveness of Cyberbullying Prevention Strategies in the UAE

Abdulla AL Nuaimi

Abstract The ever rising Internet of Things, rapid globalization and technology advancement have created a valuable seamless platform for folks to interact, publish and share contents online over high-speed connections regardless of location and distance. Nonetheless, several benefits, proliferation of vigor and anonymity of the Internet create a permanent breeding ground for all forms of cyber threats globally and the UAE in particular including all sorts of online crime like sexual harassment, deception, cyborgs, compromised safety and fraudulent engagements among others leading to manipulation of user credentials, perceptions, exposure to offensive content, financial losses, psychological and physical harm. In this paper, we examine a variety of Cyberbullying offensive and preventive strategies available globally giving particular attention to the UAE, examine the phenomena of risky Internet use of ICT tools and its relationship to Cyberbullying and assess the effectiveness of existing anti-Cyberbullying strategies used in the UAE. The findings and recommendations of the paper can be useful to government policy-maker and online crime regulators who intend to effectively curb Cyberbullying in the region.

Keywords Cyberbullying · Bullying · Cyber security · Online harassment · United Arab Emirates (UAE) · Internet of Things (IoT)

1 Introduction

In technology advancement, the rise of IoT and pivotal innovation on the World Wide Web has altered the way societies behave and interact globally on online creating a competitive atmosphere in educational, social and economic realms [3]. Meanwhile, the UAE has upgraded its cybercrime legislation to take account of technological and social change (from Federal Law No. 2/2006, to Federal Law No. 5/2012) by imposing new fences and strict penalties on information technology crimes. Bullying is normally intended to gain power over a victim especially in the case of students in

A. AL Nuaimi (✉)
SecureTech LLC, Abu Dhabi 41811, United Arab Emirates
e-mail: alneaimi@securetech.ae

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_72

primary, secondary and higher institutions of learning. Cyberbullying, on the other hand, is more severe compared to traditional bullying due to the fact that it attracts a wider scope and the bullies are anonymous. Moreover, with cyberbullying, detection may require organizations and government to recruit experts with the necessary experience to manually classify online content, whether it is harmful to the online users or not and applying unsupervised tools and learning algorithms to automatically perform the classification. Meanwhile, most Cyberbullying incidents occur away from school campuses driven by technology which allows information to spread in a short spell triggered by conflicts within schools and may lead to violence which disrupts learning and concentration of the students. Research shows that females are more emotionally offended by the inappropriate online content than males.

Additionally, the widespread use of media sites, chatrooms, temporary email accounts, Facebook, Twitter, YouTube, WhatsApp, face messenger, immo and pay as you go cell phones among others provides an avenue for bullies to remain anonymous [3] creating the need for continuous monitoring and supervision of online activities with modern technology, strong acceptable use policies and legislation and anti-bullying programmes in addition to training and awareness. Further, efforts may be required toward detection and identification of predators and online victims if we are to curb the Cyberbullying problem globally, the UAE in particular.

The rest of the paper is organized as follows: Sect. 2 provides a comprehensive review of related work and efforts from different countries regarding the effectiveness of the Cyberbullying prevention strategies applied in their countries. Section 3 reviews the Cyberbullying efforts in the perspective of the UAE as compared to other countries, especially revealing the challenges impending successful implementation of such efforts and possible solution to the challenges taking a leaf from other countries. Section 4, on the other hand, presents the paper recommendations, conclusions and future research avenues.

2 Related Work

Empirical studies and some high profile circumstantial cases globally have explained several Cyberbullying cases, challenges and prevention strategies especially in children between 5 and 10 years and the youth between 13 and 22 years due to their extensive use of the Internet and social media [2]. For instance, Abaido [1] conducted a quantitative study on the extent of Cyberbullying by collecting and analyzing data from over 200 university students in the UAE especially considering the prevalence rates of Cyberbullying, the different forms of podiums predatory to Cyberbullying, the views of the youth and whether they preferred to remain silent after different forms of Cyberbullying incidents among others. Moreover, Abaido's research revealed that social and cultural restraints among most societies in the region create the biggest impasse toward reporting and declaration of Cyberbullying incidents in university students. Further, Abaido [1] analysis revealed 91% of the students confirmed the

Table 1 Social media platforms where Cyberbullying occurs

Which of the following social media platforms have more Cyberbullying?		*N*	% age
1	Facebook	76	38.0
2	Twitter	71	35.5
3	Snapchat	31	15.5
4	YouTube	50	25.0
5	Instagram	111	55.5
6	Blogs	4	2.0

Source Abaido [1], pg. 8

existence of Cyberbullying on social media platforms with Instagram and Facebook taking the lead as shown in Table 1.

The findings from Abaido's research revealed that 72% of the respondents agreed that the Cyberbullying issue was very common among the youth between 14 and 18 years with several negative outcomes like depression, school dropouts, substance abuse and sleeping disorders among others, and the same was in turn reported by Diener West and Leaf and Privitera and Campbell [14]. Despite the previous findings deliberations and efforts narrowing down the Cyberbullying problem to the youth, the researcher envisages a deeper penetration of this problem in all age groups from 8 to over 60 years fueled by the high penetration of the Internet in the UAE to the tune of 90% by 2020, the widespread use of smart devices, the explosion of online gaming and betting, the intensification of IoT technologies and the recent outbreak of the COVID-19 global pandemic in the first quarter of 2020 that limited the public and the entire global population to working from home with the jobless limited only to social media platforms for chatting and bullying others as the only way to kill boredom, hence escalating the Cyberbullying problem globally, the UAE in particular.

Meanwhile, most of the anti-bullying approaches in schools were conducted through proactive measures like policies and supervision, reactive measures like sanctions in addition to peer support activities like school conferences, parents training, classroom ground rules, anti-bullying videos among others [11] and the reduction of known risks by protecting children online [10]. Additionally, Robertson and Amna [15] emphasized that Cyberbullying can no longer be ignored by any school management in the UAE by urging management in all schools to enforce and communicate appropriate use of Internet policies to all teachers, students and parents from time to time. Moreover, any school that fails to observe such policies shall be punished under the new Dubai schools law (Executive Council Resolution No. 2 of 2017) which requires all schools to have an obligation toward "student's safety and protection online" through due care and protection of student's rights intensified by existing telecommunication regulatory authority's Internet access management policy that sets appropriate measures, restrictions and approach through which Internet content containing pornography, nudity, gambling and any other unacceptable content to be blocked sets clear ramifications for perpetrators.

Meanwhile, Madden et al. conducted research looking into the impact of technology explosion on the lives of teenagers by examining a population of 802 teenagers aged between 12 and 17 years. Their findings revealed that 95% of the respondents had access to Internet while 78% owned a mobile phone, 47% owned a smartphone and 23% held a tablet computer implying a deep penetration into the digital world, interacting and posting all forms of personally identifiable information online leading to exposure to online crime and Cyberbullying. Moreover, the prevalent form of such incidents includes (i) harassment by sending offensive messages; (ii) sexting especially by sharing nude pictures of someone online illegally; (iii) flaming or sending vulgar messages online; (iv) cyberstalking or online threatening messages; (v) impersonation by illegally breaking into someone's account and (vi) trickery or forcing someone to reveal sensitive information online among others [9]. Figure 1 shows the most prevalent locations for Cyberbullying in children less than 10 years globally.

From Fig. 1, most of the Cyberbullying for children occurs at school and during transit in buses with a sizable figure on social media platforms which creates a major call for parents and guardians to focus on the safety of their children from anywhere any time to avoid victimization. Further, Table 2 presents the global Cyberbullying

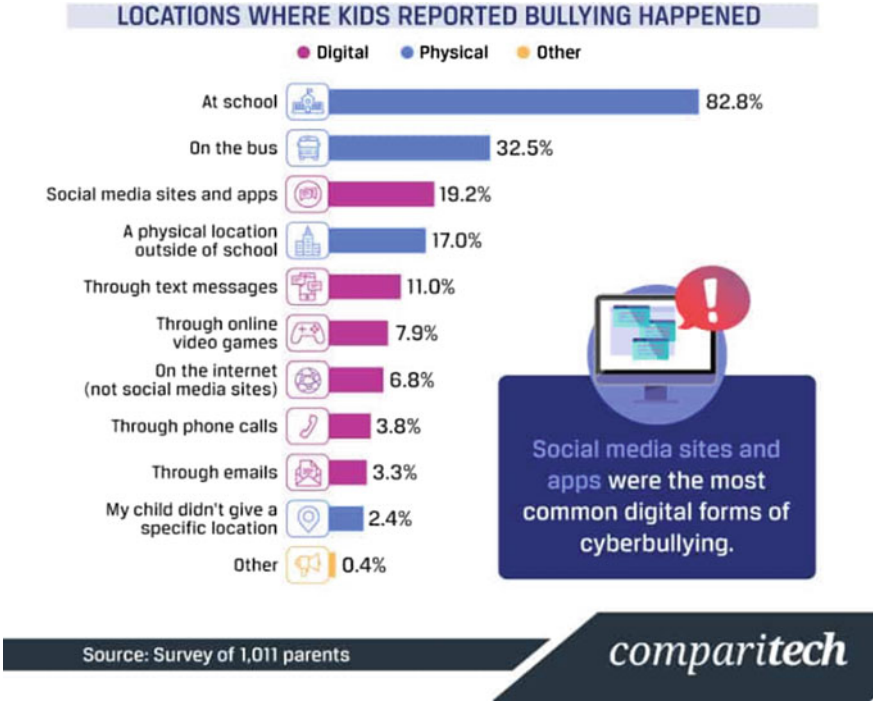


Fig. 1 Cyberbullying facts and statistics for 2020. *Source* https://www.comparitech.com/internet-providers/Cyberbullying-statistics/, 16-06-2020

Table 2 Cyberbullying statistics for the period 2011–2018

Country	Year		
	2018	2016	2011
India	37	32	32
Brazil	29	19	20
USA	26	34	15
Belgium	25	13	12
South Africa	26	25	10
Malaysia	23	0	0
Sweden	23	20	14
Canada	20	17	18
Turkey	20	14	5
Saudi Arabia	19	17	18
Australia	19	20	13
Mexico	18	20	8
Great Britain	18	15	11
China	17	20	11
Serbia	16	0	0
Germany	14	9	7
Argentina	14	10	9
Peru	14	13	0
South Korea	13	9	8
Italy	12	11	3
Poland	12	18	12
Romania	11	0	0
Hungary	10	11	7
Spain	9	10	5
France	9	7	5
Chile	8	0	0
Japan	5	7	7
Russia	1	9	5

Source Comparitech: https://www.comparitech.com/internet-providers/Cyberbullying-statistics/, visited on May 30, 2020

statistics collected over three selected years (2011, 2016 and 2018) from different countries;

Results in Table 2 show that the Cyberbullying problem has been growing over a decade across global boundaries except in Russia and Japan where very strict legislation has harnessed the impact and rise of Cyberbullying and indicated by the lowering figures from say from 9 in 2016 to 1 in 2018. Additionally, findings from the Cyberbullying research center in the USA after conducting 11 different

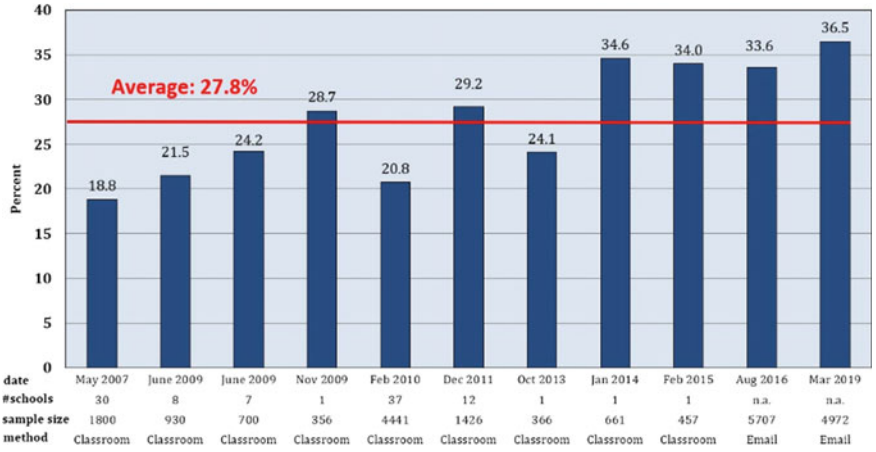


Fig. 2 Lifetime Cyberbullying victimization rates from eleven different studies, 2007–2019. *Source* Patchin and Hinduja [13]. Cyberbullying Research Center, USA

studies from 13 different projects from over 25,000 middle and high school students between 2007 and 2019 further accentuate the fact that rates of online Cyberbullying victimization have increased rapidly over the recent years. It is indicated that an average of about 27.8% students agreed to the fact that they had been victims of Cyberbullying averaged over a decade as indicated in Fig. 2.

Meanwhile, findings from Grierson indicate that every two in five children had been victims of Cyberbullying in New Zealand while three in five teenagers and women aged between 18 and 19 were victimized through online platforms. The major cause of such high statistics was attributed to the rapid increase in broadband Internet penetration and connectivity to the tune of 89.4% New Zealand and 85% Australia. Broadband Internet penetration figures are even higher in the GCC, for instance Bahrain, Qatar and the UAE are ranked among the world top ten with Internet penetration at 98.7%, 95% and 95.7%, respectively [7]. Meanwhile, most victims report poor interpersonal relations, low self-esteem and depression among others [5, 6, 12]. Moreover, bullies always take advantage of the victim's weak points like appearance, family problems and social class. Further, the need for constructs and Cyberbullying measurement tools has been called upon by several authors [4]. Batoul and others looked at the possibility of designing a multilingual system for Cyberbullying detection and assessment in Arabic language with the help of machine learning and natural language processing to support the existing English versions especially in the Arab region was profane jargons maybe stated in Arabic creating the need for filtration and classification of contents online.

3 Cyberbullying Efforts and Challenges in the UAE Perspective

The concerns of bullying and Cyberbullying have not been recently emphasized in the UAE and the entire Arabian Peninsula. This is evidenced by the few published findings accessible despite empirical evidences observed from Europe and the USA.

Meanwhile, some few scholars have attempted the phenomena over the previous decade. For instance, Kazarian and Ammar [8] studied school bullying in the Arab region through multiple literature review. Their findings show that 20.9% of middle east school juveniles report bullying victimization in UAE, 31.9% in Morocco, 33.6% Lebanon while 39.1% in Oman and 44.2% reported Cyberbullying victimization from Jordan. The authors point out the urgent need for interventions into the subject to explore signs, locations and consequences to school bullying in the region. Besides, the Dubai government established a new schools law on Cyberbullying which requires all schools to establish and implement student's online safety and protection policy as part of their cybersecurity programme (Executive Council Resolution No. 2, 2017) in addition to efforts from the UAE Telecommunications Regulatory Authority to block unauthorized sites and those with profane content in the region.

Additionally, in a survey conducted by ICDL Arabia, 60% of the youth in the GCC admitted to the presence of Cyberbullying among peers while 54% of the UAE adolescents were not aware that Cyberbullying is punishable by the existing legislation and 48% self-confessed to the use of Internet social media platforms like Instagram and Facebook without any parental supervision [7]. This claim implies cyber addiction for children which increases exposure to profane content, nudity and joining wrong groups that may persuade them into committing online crime and in turn lead to an overall change in lifestyle, poor performance in school or dropping out completely. Addiction is also fueled by the fast paced growth of cyber technology, ambient home environments with 24/7 high-speed Internet connectivity which facilitates engagement of students into off campus Cyberbullying activities and the lack of proper legislative guidance among others.

Further, Cyberbullying may lead to loss of lives if not quickly handled by specialized experts or counselors. An example is the case of Tyler Clementi of September 22, 2010, in the USA where an 18-year-old student at Rutgers University jumped off the George Washington Bridge and died after being bullied for having sexual harassment with a fellow male partner in same institution.

Therefore, in an attempt to reduce Cyberbullying in the UAE, defamation was considered a criminal offense according to (Federal Law No. 3 of 1987; Federal Decree No. 5 of 2012) in which cyber criminals and bullies are punishable by imprisonment in addition to payment of a fine at the discretion of a judge. Additionally, it is worth noting that everything done online leaves a digital footprint upon which investigation and prosecution can start or be evidenced, therefore, parents need to be vigilant to report any form of bullying incidents in the community to the nearest police station for further investigation and traceability despite the fact that victims

are always surrounded by a deep sense of fear, loneliness and helplessness with psychological and emotional scars that need the attention of experts.

4 Recommendations, Conclusions and Future Interventions

4.1 Recommendations

Based on the critical literature review and the researcher's experience in the subject matter, it is recommended that

- All schools and universities in the UAE educate their students on the dangers and legal implications of Cyberbullying, inappropriate use of the Internet and social media platforms through online safety and protection policies in addition to continuous training, awareness and social orientation programmes especially on cyber safety and risks associated with the Internet for both parents and children.
- Minimum age upon which children in the UAE are granted access to online platforms and social media be enacted by the UAE government authorities.
- Schools and universities need to conduct more investigations into the relationship between Cyberbullying and academic performance of students in addition to revealing its impact based on cross-age and cross-sex demographics.

4.2 Conclusions and Future Interventions

Much as the Cyberbullying challenge is still on the rise in many countries globally, the UAE in particular, some significant efforts have been done across the globe. Moreover, significant gaps still exist in the body of literature especially in terms of detection and identification of the victims and bullies with little efforts invested into technologies and mechanisms for uncovering accurate and reliable information on the trends and impact of Cyberbullying, limited forensic investigations and efforts into protecting the chain of custody log to ensure admissibility in courts of law. Moreover, if the integrity of evidence collected after Cyberbullying incidents is not handled by following the right forensic evidence collection and integrity protection guidelines, then most of the perpetrators will go unpunished leading to further escalation of the Cyberbullying problems globally, the UAE in particular.

Therefore, efforts are needed in designing strategies to improve digital forensic investigation, efforts to properly handle the Cyberbullying chain of custody log and tools for accurate prosecution of perpetrators by the UAE government in addition to enhancing partnerships between federal entities, private and public sector organizations, parents, schools and universities to develop consolidated efforts for identification of the bullies and victims. This will in turn address the anonymity problem, harness the jurisdictional boundary challenges through international cooperation and

collaborations, hence reducing the Cyberbullying problems in all Emirates of the UAE. In the future, the researcher intends to explore the extent of the Cyberbullying problem in the UAE public and private sector organizations.

References

1. Abaido, G. M. (2020). Cyberbullying on social media platforms among University Students in the United Arab Emirates. *International Journal of Adolescents and Youth, 25*(1), 407–420.
2. Alim, S. (2016). Cyberbullying in the world of social media: A literature review. *International Journal of Cyber Behaviour, Psychology and Learning, 6*(2), 68–95. https://doi.org/10.4018/IJCBPL.2016040105.
3. Donegan, R. (2012). Bullying and cyberbullying: History, statistics, law, prevention and analysis. *The Elon Journal of Undergraduate Research in Communications, 3*(1).
4. Haider, B., Chamoun, M., & Serhrouchni, A. (2017). A multilingual system for Cyberbullying detection: Arabic content detection using machine learning. *Advances in Science, Technology and Engineering Systems Journal, 2*(6), 275–284.
5. Hinduja, S., & Patchin, J. W. (2009). *Cyberbullying: Emotional and psychological consequences.* Cyberbullying Research Center.
6. Hinduja, S., & Patchin, J. W. (2010). *Cyberbullying by gender*. Cyberbullying Research Center. Retrieved from Cyberbullying Research Center Website http://www.Cyberbullying.us/2010_charts/Cyberbullying_gender_2010.jpg.
7. ICDL Arabia Cyber Safety Report. (2015). *Research into the online behaviour of Arab youth and the risks they face.* The International Standard in Computer Certification. icdlarabia.org.
8. Kazarian, S. S., & Ammar, J. (2013). School bullying in the Arab world: A review. *The Arab Journal of Psychiatry, 24*(1), 37–45.
9. Kraft, M. E., & Wang, J. (2009). Effectiveness of cyberbullying prevention strategies: A study on students' perspectives. *International Journal of Cyber Criminology (IJCC), 3*(2), 513–535. ISSN: 0974-2891.
10. McGuckin, C., Perren, S., Corcoran, L., Cowie, H., Dehue, F., Sevcikova, A., et al. (2013). How we can prevent Cyberbullying and how victims can cope with it. *Cyberbullying through the new media.* Florence, KY: Psychology Press.
11. Olweus, D., & Limber, S. P. (2010). Bullying in school: Evaluation and dissemination of the Olweus Bullying Prevention Program. *American Journal of Orthopsychiatry, 80*(1), 124–134. https://doi.org/10.1111/j.1939-0025.2010.01015.x.
12. Patchin, J. W., & Hinduja, S. (2006). Bullies move beyond the schoolyard: A preliminary look at Cyberbullying. *Youth Violence and Juvenile Justice, 4*(2), 148–169.
13. Patchin, J. W., & Hinduja, S. (2019). *Lifetime cyberbullying victimization rates from eleven different studies, 2007–2019*. USA: Cyberbullying Research Center. https://cyberbullying.org/summary-of-our-cyberbullying-research. June 29, 2020.
14. Privitera, C., & Campbell, M. A. (2009). Cyberbullying: The new face of workplace bullying? *Cyber Psychology & Behavior, 12*(4), 395–400.
15. Robertson, F., & Amna, Q. (2017). *Cyberbullying in the UAE: A snapshot of cyberbullying laws in the UAE* (pp. 2–3). UAE: Dubai International Financial Centre.

A Comparison Between Adaptive Neuro-fuzzy Inference System and Autoregressive Integrated Moving Average in Predicting COVID-19 Confirmed Cases in Bangladesh

Sirajum Monir Parvez, Syed Shahir Ahmed Rakin, Md. Asadut Zaman, Istiaq Ahmed, Redwanul Alam Alif, Ania-Nin-Ania, and Rashedur M. Rahman

Abstract Since December 2019, the novel coronavirus (COVID-19) has become one of the most contagious diseases to have hit the world for several decades. From December 2019 till May 2020, this respiratory syndrome-like disease has quickly spread to all countries around the world and has taken more than 400 thousand lives. The WHO declared a global pandemic situation due to the virus from March 2020. The source of this virus is not known, especially since there are no well-placed standards for its diagnosis and treatment. Several factors are involved in the spread of the disease. There have been several studies to predict or forecast the number of new cases in upcoming dates. In our study, we tested the widely used ANFIS—Adaptive Neuro-Fuzzy Inference System and the ARIMA—Autoregressive Integrated Moving Average methods to predict the total number of COVID-19 cases in the

S. M. Parvez (✉) · S. S. A. Rakin · Md. Asadut Zaman · I. Ahmed · R. A. Alif · Ania-Nin-Ania · R. M. Rahman
Department of Electrical and Computer Engineering, North South University, Plot-15, Block-B, Bashundhara Residential Area, Dhaka, Bangladesh
e-mail: sirajum.monir@northsouth.edu

S. S. A. Rakin
e-mail: syed.rakin01@northsouth.edu

Md. Asadut Zaman
e-mail: asadut.zaman@northsouth.edu

I. Ahmed
e-mail: istiaq.ahmed18@northsouth.edu

R. A. Alif
e-mail: redwanul.alif@northsouth.edu

Ania-Nin-Ania
e-mail: aina.ania@northsouth.edu

R. M. Rahman
e-mail: rashedur.rahman@northsouth.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_73

upcoming days in Bangladesh. We tuned both the models with different configuration parameters, and made 3 distinct configurations for each. After that, we applied all the different configurations on the same dataset, and the results were compared against each other in terms of statistical performance measures such as Mean Absolute Percentage Error (MAPE), Root Mean Squared Relative Error (RMSRE), Root Mean Squared Relative Error (RMSRE).

Keywords COVID-19 · Adaptive neuro-fuzzy inference system (ANFIS) · Autoregressive integrated moving average (ARIMA) · Forecasting

1 Introduction

Coronavirus, a SAARS COV-2 like virus is responsible for causing a contagious flue like disease. The disease is named COVID-19 and has already created a global pandemic situation. First discovered in the Hubei city of the province Wuhan in China, this viral disease has now reached across every country around the world. As governments and health agencies struggle to prevent the spread of coronaviruses, they need every help. As no confirmed medication has been proposed by the WHO for COVID-19 yet, the most effective way to fight against the pandemic is to prevent it from spreading by predicting possible future cases and taking precautionary measures accordingly. In recent months, the trend and final dimension of the COVID-2019 pandemic have been forecasted by an increasing number of research works using various methods. Zhao et al. [13] in their work forecasted confirmed cases in China, and tried to estimate the number of unreported cases from Wuhan. The authors found the number to be at least 469 cases from 1 to 15 January 2020. Besides, this evidence had increased 21 folds after 17 January 2020. Al-qaness et al. [1] used the ANFIS prediction tool combined with Flower Pollination Algorithm (FPA) and the Salp Swarm Algorithm (SSA) to predict confirmed cases in China. Ardabili et al. [2] used stand-alone ANFIS to predict the cases. On the other hand, Perone et al. [10], Benvenuto et al. [4], Dehesh et al. [6], etc. adopted ARIMA for their prediction and observed good results. Kumar et al. [7] used ARIMA to predict the top 15 countries by April 2020 in terms of COVID infection. Pal et al. [9] used a Bayesian optimization framework to predict risk category of a country, where the proposed a shallow long short-term memory (LSTM) based neura network.

In countries like Bangladesh, the fight against the VIRUS is even more difficult concerning the poor infrastructure and the scarcity of tools. In our study, we vowed on finding a forecasting tool that will be best suited for predicting the number of COVID patients in Bangladesh. This data will help the concerned authority to take adequate precautionary measures to stall the spread of the infection. We have applied ANFIS and ARIMA to predict confirmed cases for upcoming days using Bangladesh's COVID dataset for confirmed cases [12]. We have implemented the ANFIS model with three separate configurations for the prediction, and also made use of the ARIMA model with three different configurations to be applied on the

same dataset. Later we have compared output from all the different configurations based on a few performance parameters and proposed the best method out of them, which can be effectively utilized.

2 Methods

2.1 Fuzzy Inference System and ANFIS

Fuzzy Inference Systems take inputs and process them based on the prespecified rules to produce the outputs. Both the inputs and outputs are real-valued, whereas internal processing is based on fuzzy rules and fuzzy arithmetic. An ANFIS is a mixture of an adaptive neural network (ANN) and a fuzzy inference system. It is being used by many scholars due to its rapid learning capacity and the ability to capture the nonlinear structure of a process. Aside being used to forecast COVID cases recently, ANFIS has numerous applications in the past for other predictions as well. For instance, Mohaddes et al. [8] used ANFIS to forecast Iran's agricultural product export, Zheng et al. [14] in their work used for short-term wind power prediction. Benmouiza et al. [3] used ANFIS with subtractive clustering and greed partitioning for an hour-ahead solar radiation forecasting. ANFIS architecture uses both artificial neural networks and fuzzy logic. The parameters of the fuzzy inference system are determined by the neural network. ANFIS can approximate to any degree of accuracy any real continuous function of a compact set of parameters. Since the system is based on a fuzzy inference system which reflects incredible information, it should always be translated into fluid IF-THEN rules.

- Rules—the if-then rules have to be determined somehow. This is mostly done by 'knowledge acquisition' from an expert. It is a time-consuming process that is fraught with problems.
- Membership Functions: Researchers and Data Scientists use membership functions for determining a full-fledged fuzzy set. For the case of Gaussian functions, there should be parameters set into the Gaussian function.

ANFIS stands for Adaptive Neuro-Fuzzy Inference System; as the name suggests, it is an adaptive network which has a structured network of nodes and directional links. ANFIS involves the use of a database that contains the required rules and membership functions for learning. The adaptive network of ANFIS consists of a set of rules for learning the data, such as backpropagation. As the name suggests, it is called adaptive for having parameters that could potentially bring alterations to the node of the output. Through the networks, a relationship between input nodes and output nodes can be established and further evaluated. These adaptive networks used in ANFIS have many ways to be implemented for different and unique purposes. In our case, we will be using the method initially brought up by Jang, which is the ANFIS method shown below. The fixed nodes in the diagram are symbolized by

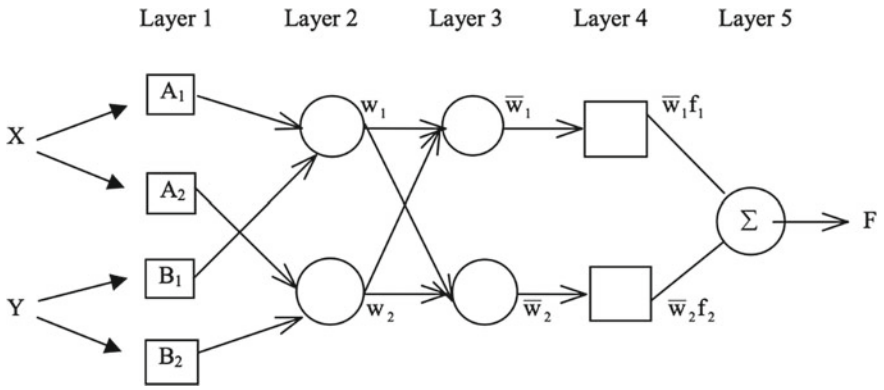


Fig. 1 ANFIS architecture based on a two rule Sugeno system

circular nodes, while the learned parameters are given shape through square nodes (Fig. 1).

$$\text{If } x \text{ is } A_i \text{ and } y \text{ is } B_i \text{ THEN } f = p_i x + q_i y + r_i$$
$$\text{If } x \text{ is } A_{i+1} \text{ and } y \text{ is } B_{i+1} \text{ THEN } f = p_{i+1} x + q_{i+1} y + r_i$$

There is a forward and a backward pass for the training of the network. We look at the forward transfer on each layer in turn. The vector input is spread by network layer by layer through the forward pass. The error is returned in the backward transfer to the backpropagation through the network in the same way. It can be anything to membership. For instance, the following function represents a Gaussian membership function.

$$\mu(x) = e^{-\left(\frac{x-\rho_i}{\alpha_i}\right)^2} \tag{1}$$

where the generalized Gaussian membership functions denoted by μ, A_i and B_i define the membership values of μ. a_i, r_i denotes the premise parameters set.

The result/output of each layer is fed to the next layer as input with a weight assigned to it and in the final layer summation of all the nodes outputs are generated as final output, and is represented by

$$\sum \bar{\omega}_i f_i \tag{2}$$

2.2 Autoregressive Integrated Moving Average (ANFIS)

Autoregressive Integrated Moving Average (ARIMA) is a prediction method that projects future values of a series based entirely on their inertia. It primarily uses short-term projections, which requires at least 40 historical data points. This works very well if the data display a steady or stable trend with a minimum of outliers over time. Often referred to as Box-Jenkins method at the time when the data are relatively long and the association between past observations are stable, ARIMA is generally superior to exponential smoothing techniques. ARIMA is preferred for prediction using time-series data and it performs best when there is seasonality in the dataset. ARIMA has been used by many to predict time series data, for example, in 2018 Wadi et al. [11] used Amman Stock Exchange's 8 years data from 2010 to 2018 for colsed time series prediction. Rebane et al. [15] used ARIMA for cryptocurrency price prediction and compared the results of the model with a deep multi-layer Seq 2 Seq RNN model.

ARIMA models are always expressed with the help of a few parameters, and the model is expressed as ARIMA (p, d, q). In this case, p means the order of self-regression, d means the degree of trend variance, and q means the average of movements. The autocorrelation function (ACF) graph and partial autocorrelation (PACF) graph is utilized to find the initial number of ARIMA models. The difference in normality and stationery is then evaluated for ARIMA models. After that, they are checked for accuracy by observing their MAPE, MAE, and RMSE values to determine the finest model to forecast. The model for forecasting the number of future confirmed COVID-19 cases is represented as,

$$\text{ARIMA}(p, d, f)\colon X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + Z_t \tag{3}$$

$$\text{where}\ , \quad Z_t = X_t - X_{t-1} \tag{4}$$

Here, X_t is the predicted number of confirmed COVID-19 cases at tth day, α_1, α_2, β_1, and β_2 are parameters whereas Z_t is the residual term for tth day. In previous cases, the pattern of potential effects can be predicted and a time-series study is performed to that end.

3 Experiment

This section presents the description of the used dataset, the performance measures, the parameter setting for all methods, the experiment results, and discussions.

3.1 Data Description

The main dataset of this study is extracted from the Worldometers.com (https://www.worldometers.info/coronavirus/country/bangladesh/), a reference website providing counters and real-time statistics on diverse topics. It is owned by data company Dadax Limited [12]. It contains the daily confirmed cases, deaths, and recovery reports in Bangladesh from 16 March to date, as shown in Table 1. The COVID 19 crisis

Table 1 Bangladesh's COVID-19 dataset

Date (D/M/Y)	New case	Total case	Date (D/M/Y)	New case	Total case	Date (D/M/Y)	New case	Total case
16/3/20	3	3	14/4/20	209	1005	13/5/20	1162	17,795
17/3/20	0	3	15/4/20	219	1224	14/5/20	1041	18,836
18/3/20	4	7	16/4/20	341	1565	15/5/20	1202	20,038
19/3/20	4	11	17/4/20	266	1831	16/5/20	930	20,968
20/3/20	2	13	18/4/20	306	2137	17/5/20	1273	22,241
21/3/20	4	17	19/4/20	312	2449	18/5/20	1602	23,843
22/3/20	3	20	20/4/20	492	2941	19/5/20	1251	25,094
23/3/20	6	26	21/4/20	434	3375	20/5/20	1617	26,711
24/3/20	6	32	22/4/20	390	3765	21/5/20	1773	28,484
25/3/20	0	32	23/4/20	414	4179	22/5/20	1694	30,178
26/3/20	5	37	24/4/20	503	4682	23/5/20	1873	32,051
27/3/20	4	41	25/4/20	309	4991	24/5/20	1532	33,583
28/3/20	0	41	26/4/20	418	5409	25/5/20	1975	35,558
29/3/20	0	41	27/4/20	497	5906	26/5/20	1166	36,724
30/3/20	1	42	28/4/20	549	6455	27/5/20	1541	38,265
31/3/20	2	44	29/4/20	641	7096	28/5/20	2029	40,294
1/4/20	3	47	30/4/20	564	7660	29/5/20	2523	42,817
2/4/20	2	49	1/5/20	571	8231	30/5/20	1764	44,581
3/4/20	5	54	2/5/20	552	8783	31/5/20	2545	47,126
4/4/20	9	63	3/5/20	665	9448	1/6/20	2381	49,507
5/4/20	18	81	4/5/20	668	10,116	2/6/20	2911	52,418
6/4/20	35	116	5/5/20	786	10,902	3/6/20	2695	55,113
7/4/20	41	157	6/5/20	790	11,692	4/6/20	2423	57,536
8/4/20	54	211	7/5/20	706	12,398	5/6/20	2828	60,364
9/4/20	112	323	8/5/20	709	13,107	6/6/20	2635	62,999
10/4/20	94	417	9/5/20	636	13,743	7/6/20	2743	65,742
11/4/20	58	475	10/5/20	887	14,630	8/6/20	2735	68,477
12/4/20	139	614	11/5/20	1034	15,664	9/6/20	3171	71,648
13/4/20	182	796	12/5/20	969	16,633			

definitely has increased the visibility of the website. It is one of Google's highest-ranking coronavirus search results. Such time-series data are collected by the State, local governments, and health authorities from monitoring the ongoing epidemic.

Worldometres.com has been monitoring coronavirus cases in real-time since late January, as they have been identified following research. However, due to the broad lack of testing, the data in the outbreak picture is necessarily limited. We used data from 16 March 2020 to 30 May, 2020, to train the model. Data from 31 May 2020 to 9 June 2020 is used to test. The death cases were not considered as they do not have any relevance in the forecasting for confirmed cases. Therefore, data have been filtered to remove the death case column. Then we have done comparative analysis between ANFIS and ARIMA based on cumulative confirmed case forecasting. The key indicators for determining the quality of performance are the root mean squared error. Mean absolute error and mean percentage error.

3.2 Performance Measure and Parameter Settings

The quality of the proposed method is evaluated using a set of performance metrics as follows:

- Root Mean Square Error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_s}\sum_{i=1}^{N_s}(\mathrm{YYP}_i - Y)^2} \tag{5}$$

- Mean Absolute Error (MAE)

$$\mathrm{MAE} = \frac{1}{N_s}\sum_{i=1}^{N_s}|\mathrm{YYP}_i - Y_i| \tag{6}$$

- Mean Absolute Percentage Error (MAPE)

$$\mathrm{MAPE} = \frac{1}{N_s}\sum_{i=1}^{N_s}\left|\frac{\mathrm{YP}_i - Y_i}{\mathrm{YP_i}}\right| \tag{7}$$

We used ANFIS and ARIMA to predict the outcomes from the data set. We tuned the ANFIS model with different parameter settings to bring variation in the results, so did we do with the ARIMA model as well. We prepared 3 ANFIS models and

Table 2 Parameter settings for the three separate ANFIS configurations

Settings	ANFIS-1	ANFIS-2	ANFIS-3
Cluster type	Grid partitioning	Grid partitioning	Subtractive clustering
Input	Gaussian MF	Generalized bell-shaped MF	0.7
Output	Linear MF	Linear MF	0.3
No. of membership functions	3	5	–
Epochs	50	100	50
Step size	1.10	1.10	1.30

3 ARIMA models to with minor changes in the parameters. Then we computed the result to determine which model from the ANFIS and the ARIMA performed best individually. After that the best models from both cases were analyzed to observe the performance difference between ANFIS and ARIMA. Table 2 shows the parameter settings for the 3 distinct ANFIS models.

We first created an ANFIS model with 5 Gaussian membership functions and the input space was divided by Grid partitioning. The step size increase rate was set to its default 1.10 and the dataset was trained in 50 full cycles. For ease of understanding, we called this configuration to be ANFIS-1. The input membership function was then tuned and was changed to Generalized Bell-shaped MF for the second configuration, which was ANFIS-2. In this case the total number of membership functions were 5, and the Epoch was set to 100, while the step increase size remained the same. Figure 2a and b represents input membership functions for both these configurations.

For the third configuration, we went for a different clustering option by choosing subtractive clustering. The configuration is mentioned in Table 2.

On the other hand, for the ARIMA model we created 3 partitions by changing only the value for order of differencing, d. Its value for ARIMA-1 was set to 1 and for ARIMA-2, and for ARIMA-3, the value was set to 2 considering the data to be seasonal in nature this time The other parameters remained the same. The parameter configurations are depicted in Table 3.

When we analyze the data to extract insights, we find that our ANFIS-3 configuration has the lowest percentage of error. It obtains Mean Absolute Percentage Error value of 2.46%. Among the ARIMA configurations, our ARIMA-3 model earns a value of 3.26%.

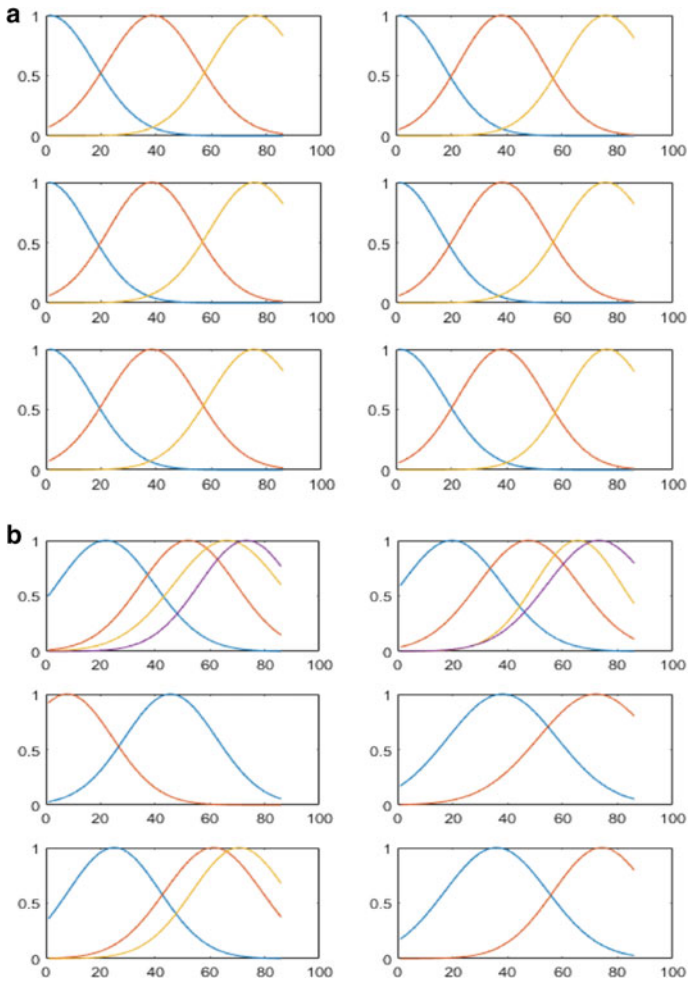


Fig. 2 **a** Input membership function ANFIS-1 configuration. **b** Input membership function ANFIS-2 configuration

4 Results and Analysis

All 3 ANFIS models and 3 ARIMA models performed quite well in prediction. The results with respect to the original testing value is given in Table 4.

After analyzing the forecasted values, it is found that the ANFIS-3 configuration has the lowest percentage of error. It obtained Mean Absolute Percentage Error value of 2.46%. Among the three ARIMA configurations, the ARIMA-3 model earned a value of 3.26%. Performance parameters of the above results are mentioned in Table 5.

Table 3 Three separate ARIMA models' parameter settings

Parameter	ARIMA-1	ARIMA-2	ARIMA-3
The order (no. of time lags) of the autoregressive ("AR") model	1	1	2
The order of the moving average ("MA") model	1	1	2
The maximum value of p	3	4	5
The maximum value of q	3	4	5
Period of seasonal differencing	3	3	3
The order of first-differencing	1	2	2
The order of seasonal differencing	1	1	1

Table 4 Forecasted results from all the configurations

Date (D/M/Y)	Actual cases	ANFIS -1	ANFIS-2	ANFIS-3	ARIMA-1	ARIMA-2	ARIMA-3
31/5/20	47,126	45,684	46,008	46,438	46,526	46,574	46,655
1/6/20	49,507	47,429	47,791	48,385	48,466	48,595	48,807
2/6/20	52,418	49,173	49,552	50,331	50,401	50,642	51,000
3/6/20	55,113	50,917	51,276	52,266	52,330	52,717	53,267
4/6/20	57,536	52,659	52,948	54,184	54,254	54,819	55,615
5/6/20	60,364	54,399	54,555	56,074	56,173	56,948	58,007
6/6/20	62,999	56,135	56,082	57,929	58,086	59,105	60,476
7/6/20	65,742	57,869	57,516	59,740	59,994	61,289	63,029
8/6/20	68,477	59,600	58,849	61,498	61,898	63,499	65,628
9/6/20	71,648	61,328	60,073	63,193	63,796	65,738	68,308

Table 5 Comparison of all ANFIS and ARIMA configurations based on the performance parameters

Method		RMSE	MAE	MAPE (%)
ANFIS	ANFIS-1	6231.39	5573.7	8.96
	ANFIS-2	6512.42	5628.0	8.95
	ANFIS-3	4751.37	1484.5	2.46
ARIMA	ARIMA-1	4513.39	3900.6	6.20
	ARIMA-2	3518.29	3100.4	4.96
	ARIMA-3	2199.41	2013.8	3.26

Table 5 shows a relative comparison between the ANFIS and ARIMA models performance parameters. It shows both model's performance measures in terms of Root Mean Square Error, Mean Absolute Error, Mean Absolute Percentage Error. If we look closely, we see that out of the three configurations we did for the ARIMA

models, all of them performed way better than ANFIS-1 and ANFIS-2, while ANFIS-3 was the best among all the six. ANFIS-1 and ANFIS 2 both used Grid partitioning for input spacing. On the other hand, ANFIS-3 used Subtractive clustering technique. This clearly indicates the high-level of performance due to the change in clustering type or partitioning choice. Though ANFIS-2 utilized 100 epochs, a different membership function with a higher number of membership functions, its performance could not be updated significantly. In fact, it's only 0.01% improved than ANFIS-3. Meanwhile all the ARIMA models performed well. With the increase in the number of orders of differencing, the performance went well. With the addition of seasonality consideration in the data, the performance improved significantly for the ARIMA-3 configurations.

Figure 3a and b represents the ANFIS 3 configurations' prediction curve, and the increase in the percentage of cumulative cases. Figure 3a depicts graph of 10 days'

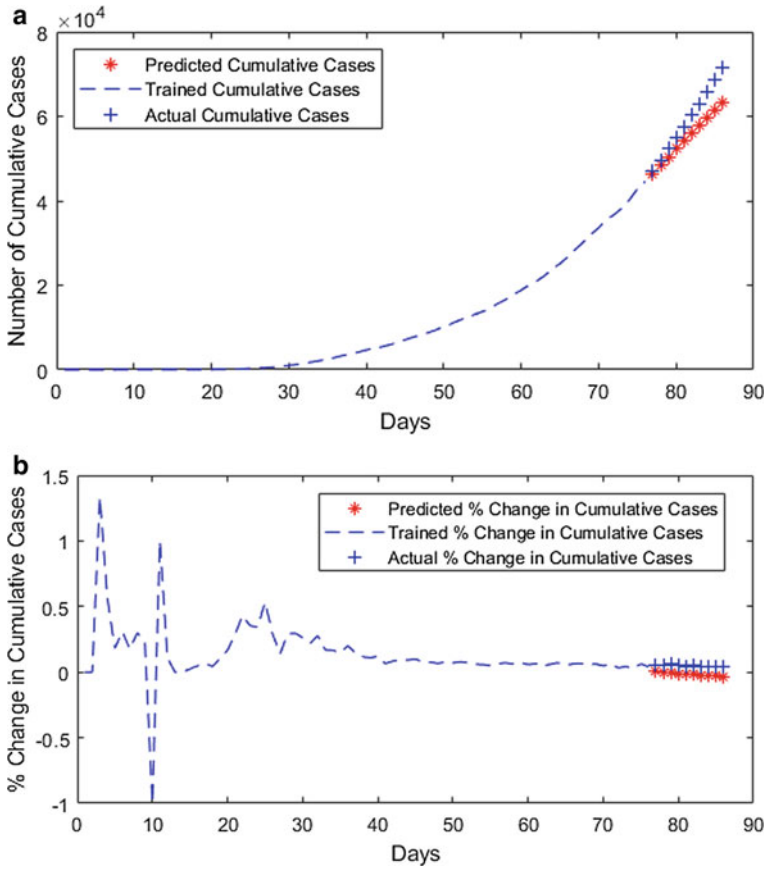


Fig. 3 **a** Graph representing prediction by ANFIS-3 configuration. Blue dashed line depicts trained data, and blue cross denoted part indicates tested data, while the red stars show predicted values. **b** Percentage increase in cases from ANFIS-3 model

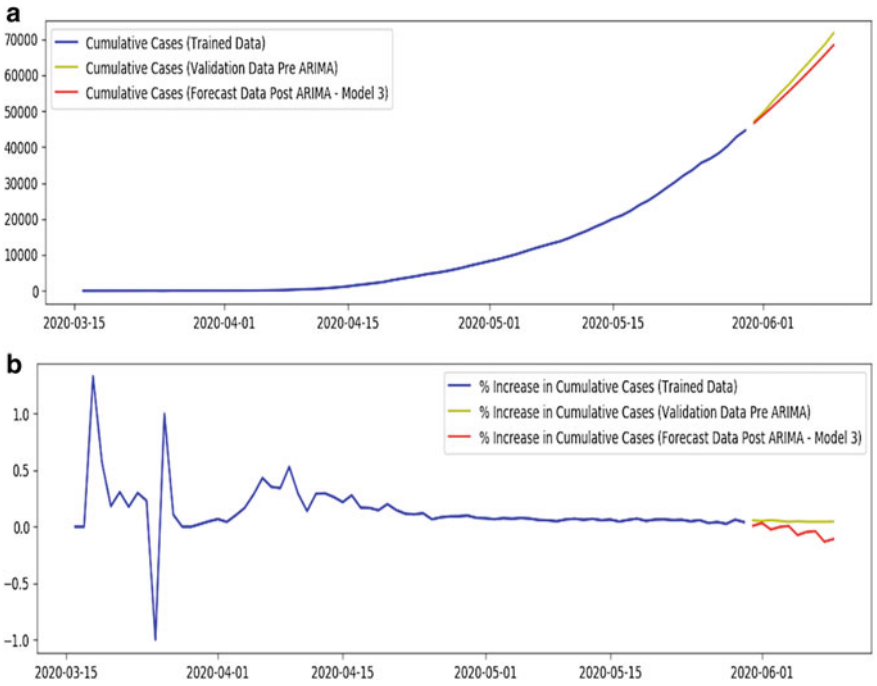


Fig. 4 **a** ARIMA 3 configuration. The blue line indicates cases in the trained data, while the amber line shows the validation data for pre-ARIMA and the red line shows the forecasted data for post-ARIMA. **b** Representation of the percentage increase in cases from ARIMA-3 model

(from 31 May till 09 June) prediction in terms of number of cumulative cases. On the other hand, Fig. 4a and b represents the graph of the same indicators produced by the ARIMA 3 configuration.

5 Future Scope of Work

For Prediction of COVID-19 Cases in Bangladesh, we can also use Holt-Winters Exponential Smoothing. It is another suitable forecasting model data scientists use to predict values such as stock market values and other factors that change on a day-to-day basis. The Holt-Winters Exponential Smoothing is also called Triple Exponential Smoothing, as it adds seasonality factor to the existing time series. It has two types of seasonality: Additive and Multiplicative. Additive Seasonality is used in the case of linear exponentiality and Multiplicative Seasonality is used for exponential seasonality [5]. More data for the COVID-19 cases will be added concerning Bangladesh for further analysis as there have been cases every day in this country. Through the addition of data, we can further analyze COVID-19 cases in

Bangladesh. Along with the new addition of data, we can improve the results through ARIMA, ANFIS, and Holt-Winters Smoothing.

6 Conclusion

This work focused on utilizing two widely used forecasting methods, the Adaptive Neuro-Fuzzy Inference System and the Autoregressive Moving Average method, to see their performance in predicting COVID 19 cases in Bangladesh. This paper also compared both the models with respect to a set performance measures and tried to identify the best-suited method for the said prediction. Both of the models have used cumulative cases and percentage change in cumulative cases against dates. ANFIS and ARIMA was tuned with 3 different parameter settings. ANFIS with subtractive clustering configuration performed the best and showed the least amount of error in test data. For ARIMA, we had to start from the basic required parameters to see how the predictions could turn out. Before training the data in both models, some manual effort was involved for calculating percentage change in cumulative COVID-19 cases in Bangladesh. To summarize, it has been found that both the models are good in prediction but requires further improvement. Other coexisting algorithms can be studied and included with each of the methods mentioned above to optimize their performance further.

References

1. Al-qaness, M. A., Ewees, A. A., Fan, H., & Abd El Aziz, M. (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. *Journal of Clinical Medicine, 9*(3), 674.
2. Ardabili, S. F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A. R., Reuter, U., Rabczuk, T., Atkinson, P. M. (2020, Apr 19). Covid-19 outbreak prediction with machine learning. Available from SSRN 3580188.
3. Benmouiza, K., & Cheknane, A. (2019). Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. *Theoretical and Applied Climatology, 137*(1–2), 31–43.
4. Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in brief, 29*, 105340.
5. Brownlee, J. (2020). A gentle introduction to exponential smoothing for time series forecasting in Python. Available from https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/. Accessed June 21, 2020.
6. Dehesh, T., Mardani-Fard, H. A., & Dehesh, P. (2020, Jan 1). Forecasting of COVID-19 confirmed cases in different countries with ARIMA models. medRxiv.
7. Kumar, P., Kalita, H., Patairiya, S., Sharma, Y. D., Nanda, C., Rani, M., et al. (2020, Jan 1). Forecasting the dynamics of COVID-19 pandemic in top 15 countries in April 2020: ARIMA model with machine learning approach. medRxiv.
8. Mohaddes, S. A., & Fahimifard, S. M. (2015). Application of adaptive neuro-fuzzy inference system (ANFIS) in forecasting agricultural products export revenues (case of Iran's agriculture sector). *Journal of Agricultural Science and Technology, 17,* 1–10.

9. Pal, R., Sekh, A. A., Kar, S., & Prasad, D. K. (2020, Mar 31). Neural network based country wise risk prediction of COVID-19. arXiv preprint arXiv:2004.00959.
10. Perone, G. (2020, Jan 1). An ARIMA model to forecast the spread and the final size of COVID-2019 epidemic in Italy. medRxiv.
11. Wadi, S. A., Almasarweh, M., Alsaraireh, A. A., & Aqaba, J. (2018). Predicting closed price time series data using ARIMA model. *Modern Applied Science, 12*(11).
12. Worldometer. (n.d.). *Covid 19 situation report Bangladesh.* Retrieved from www.worldometers.info.
13. Zhao, S., Musa, S. S., Lin, Q., Ran, J., Yang, G., Wang, W., et al. (2020). Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: A data-driven modelling analysis of the early outbreak. *Journal of Clinical Medicine, 9,* 388.
14. Zheng, D., Eseye, A. T., Zhang, J., & Li, H. (2017). Short-term wind power forecasting using a double-stage hierarchical ANFIS approach for energy management in microgrids. *Protection and Control of Modern Power Systems, 2*(1), 13.
15. Rebane, J., Karlsson, I., Denic, S., & Papapetrou, P. (2018). Seq2Seq RNNs and ARIMA models for cryptocurrency prediction: A comparative study. SIGKDD Fintech, 18, 2–6

Sensor Searching Techniques in Internet of Things: A Survey, Taxonomy, and Challenges

Farzana Noshin Choudhury, Fahrin Rahman, Md. Rafsan Jamil, and Nafees Mansoor

Abstract Along with the fast development of sensing technologies and smart devices, endless automation opportunities are expected in every sphere of modern life. Hence, the IoT concept plays a vital role in managing and controlling the wireless devices over the Internet. Therefore, IoT is considered to be the third wave of information and communication technology (ICT) after eras of the Internet and cellular networks. Besides, the post-COVID-19 world is expected to more sensor-centric to ensure lesser human interactions. Since a humongous number of sensors are intended to be deployed in IoT, efficient data communication requires the network to be clustered physically or logically. Consequently, the selection of the appropriate sensor(s) for data processing and gathering is vital in IoT. Several sensor selection techniques in IoT have been proposed recently, still, sensor searching remains quite a new research field. Therefore, several sensor searching methods for IoT are studied and presented in this paper. Moreover, the strengths and limitations of the existing sensor searching techniques are also outlined in the paper. Hence, the new performance metrics are presented in the paper, where the existing techniques for searching are analyzed.

Keywords Internet of things · Routing protocols · Sensor searching algorithm · IoT survey · Sustainable cities and communities

F. N. Choudhury · Md. R. Jamil
Ahsanullah University of Science & Technology, Dhaka, Bangladesh
e-mail: noshin.eee@gmail.com

Md. R. Jamil
e-mail: rafsan.shovon@gmail.com

F. Rahman
Islamic University of Technology, Dhaka, Bangladesh
e-mail: fahrin_rahman@yahoo.com

N. Mansoor (✉)
University of Liberal Arts Bangladesh (ULAB), Dhaka, Bangladesh
e-mail: nafees@ieee.org

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_74

1 Introduction

Internet of things (IoT), a pervasive technology, is getting noticed in the communication research circle for the past few years [1]. The term "IoT" came into the spotlight after Kevin Ashton coined it into 1999. IoT is the wide network containing physical objects or "Things" planted with electronics, software, and connectivity. The objective of IoT can be described as enabling and achieving bigger value and service through exchange of data with the manufacturer, operator, and/or other connected devices. Exchanging information or data is based on the infrastructure of the International Telecommunication Union's Global Standards Initiative [2]. The "Things" in the IoT can be considered as an animal having a biochip transponder or a device with a built-in sensor. Hence, these sensors are implemented around various geographical locations. Those sensors are also managed by different organizations.

Search engines are considered as the most vital part of the Internet. People need search engines to find their desired information in the shortest possible time. Searching in IoT is more challenging in case of searching for documents on the Web using search engines. Complex descriptions can be a feature of IoT. Thematic, spatial, and temporal dimensions can be the mode of representation of the data produced by IoT objects [3]. Again, loss of communication or malfunctioned wireless sensor nodes can bring change in the status of IoT objects. So, traditional searching techniques are not appropriate and sufficient for IoT-based applications.

Sensors implemented around the world are increasing day by day. The observation and maintenance (O&M) data of the sensors are of high importance in the Internet of things. The research and survey considering sensor searching techniques in IoT are not sufficient so far [4]. The existing searching techniques in IoT, their advantages and drawbacks, and their comparative performance analysis have been presented here.

Organization of the paper can be mentioned as the following: In Sect. 2, a discussion on different recently developed sensor searching techniques for IoT is presented. The proposed performance metrics are discussed in Sect. 3. An analysis of the existing techniques based on the proposed metrics is described in Sect. 4. Conclusion and future works have been discussed in Sect. 5.

2 Existing Techniques

In recent years, different searching techniques have been introduced to the field of IoT. A large volume of data is generated from numerous sensors. Appropriate searching technique matters a lot in collection and processing of that data. This section represents a brief discussion on the existing techniques.

Context-aware sensor search, selection and ranking model, abbreviated as CASSARAM, is a user-priority-based sensor searching technique in IoT [5]. In CASSARAM, requirements of the users are divided into the following categories:

point-based and proximity-based. In this algorithm of CASSARAM, point-based requirements that are identified and specified by the user are non-negotiable entities in the CASSARAM technology. Moreover, in CASSARAM, proximity-based requirements are considered to be negotiable where these requirements are not always considered to be user-specific. Proximity-based requirements include characteristics of sensors that include reliability, accuracy, etc., which refer to context property. Semantic Sensor Network Ontology (SSNO) is the modeling technique of sensor descriptions and context properties. SSNO stores the context properties in original measurement units (accuracy in percentage or latency). For consistency, the context properties are normalized to [0, 1]. Moreover, relational expression-based filtering is used to reduce the complexity of large queries. Finally, CPWI decides the rank of the sensor on proximity-based user requirements. To overcome the limitation of CASSARAM, comparative priority-based heuristic filtering (CPHF) is applied which performs the function of removing the sensors positioned far away from the user-defined ideal sensors [5]. In this case, distributed sensor searching can improve scalability and efficiency.

Searching with textual metadata does not work well in practice. Because search input can be in different forms. When we enter data, there are human-based mistakes. Different terms are used to describe the similar concept or important metadata is not entered. Adopting the traditional search-by-example approach can solve this problem to some extent [6]. A time series of sensor values are specified by the user for performing specific search. A comparison is made between the time series and the indexed fuzzy sets. A fuzzy set is a mapping of a set of real numbers that allows partial membership instead of binary membership. It is computed by the sensor itself. Fuzzy sets-based method can be used to compute a similarity score for each indexed sensor based on their output. When the similarity score is higher, then the measurement ranges overlap more, and more the sensor measurements belong to the fuzzy set defined by the output of the given sensor [6].

Numerous sensors are functioned for Internet access purpose. But there are some problems to ensure efficient searching. User has to model sensor data as Bayesian network (BN). This BN model is essential for forming perfect prediction of desired sensors and should be constructed, and then the probability table can be formed. BN is such a model by which a small probability table can be formed at every node which have all nodes connected to the network. Correlations can be modeled as the dependency relationship between pair of nodes. Sought state, number of matching sensors, total set of sensors, and set of sensors (contracted) are the parameters here. Bayesian network table is also used here to separate the sensors which are non-contacted [7].

For network modeling, distributed regression is used which is an efficient and general framework. The algorithm follows kernel linear regression. Shape, size of sensor data, and actual complete sensor data can be extracted through this research method [8]. The complexity and requirements can be limited so that the regression method can work perfectly to express the perfect structure and type of data. Thus, it is reducing the cost. There are two layers—junction tree and routing tree, and they are for probability function and data collection. From this model, the particular

data or model coefficient can be transmitted. This is a full algorithmic work where sensor functions and basis matrices can be implemented to get dot product matrix and project measurement vector. From these, particular coefficient can be found. The routing tree can perform transmission. It is also making it possible to access new locations. Distributed regression algorithm is such an algorithm that is built on a distributed application of Gaussian elimination. The complexity of this algorithm depends on the cluster size of each node but does not depend on the number of parameters. This research method can make connection between two sensor nodes even though they are not neighbors. Error correction can be performed at the time of implementation [8].

Shodan is a powerful IP search engine which enables notification for the IP exposure for the IoT devices. This web-based notification system is developed because the techniques for preventing IP-related security threats have been proven insufficient. Searching data for IP addresses is provided by Shodan. In just one month, Shodan engine collects approximate 500 + million IoT devices [2]. Computers that are connected to the Internet are found by the user by means of various filters in Shodan. It provides two types of information that is Web page and script console using IP search engine application programming interface (API). The Web page provides only the basic data (IP address, connected country, city etc.) and also a server name. But script console contains more useful data like port number, affiliation, latitude, longitude, etc., which are inherently more useful than data from Web page. The script API of Shodan is used to set up the IP exposure notification system. In this system, exposed IoT device information is objectified. Object is marked on Google Maps from longitude and latitude information. From these marks, it can be identified how many IP exposures are there in that particular area. If advantage taken, dangerous IP exposure can be denoted.

3 Performance Metrics

To get a bird's eye view detail of the research about existing searching techniques, it is convenient in defining some parameters or dimensions against which the existing techniques can be compared. The metrics are presented as follows:

- User's criterion: A technique for providing the opportunity to input user's criterion that can generate output matching user's requirements. A user-friendly technique consumes less time to perform the searching in IoT as the user gets easily acquainted with the method.
- Realistic application: It describes whether the working prototype supports experiments using realistic applications or not. Experiments using realistic applications demonstrate how much the technique is effective in a realistic environment.
- Provision for sensing as a service: Sensing-as-a-service model is one of the greatest priorities in IoT today. It refers to the new Internet of things business model. Till now infrastructure as a service (IaaS), software as a service (SaaS),

etc., have been introduced in IoT. Provision for sensing as a service enables pay-per-use system and free and paid system for sharing.

- Scalability: Scalability means the capacity of handling a growing amount of work by a system, network, process, or application. A scalable searching technique can deal with millions of sensor nodes which is an advantage for IoT.
- Robustness: Millions of "Things" are connected to IoT. So, the searching technique needs to be robust as any kind of failure in the sensor networks affects the data collection process. Robust techniques encourage effective searching in IoT for a long time period without disruption.
- Machine learning: Machine learning refers to artificial intelligence. In this process, computers become able to learn without being explicitly programmed. IoT devices are generating a large amount of data. In this case, machine learning can help to improve efficiency and reduce costs by pursuing those data.
- Access to the new location: In a particular technique, a sensor network should identify the new location. It should have the ability to access various locations where sensors are deployed.
- Threat detection and prevention: Security is a very important issue in IoT because all the devices are connected through the Internet. Any kind of harm can reach the devices as IP-related threats are increasing. Hence, threat detection and prevention are essential.

4 Performance Analysis

This section represents the overall comparative analysis of the techniques. Basically, the analysis can help the user to select the best searching method satisfying his requirements (Table 1).

In CASSARAM, context information is used to select sensors matching user's criteria. The algorithm used in this technique can capture user priorities.

Table 1 Summary of analysis

Technique	User's criterion	Sensing as a service	Scalability	Machine learning	Access to location	Threat detection
CASSARAM	Yes	Yes	Not enhanced	Supported	No	No
Fuzzy sets	Yes	No	Enhanced	Not supported	Yes	No
Bayesian network	Yes	No	Enhanced	Not supported	Yes	No
Distributed regression	Yes	No	Enhanced	Not supported	Yes	No
Shodan	Yes	No	Not enhanced	Not supported	No	Yes

CASSARAM also supports realistic application and is helpful for broadening Sensing-as-a-service vision. Scalability along with efficiency is enhanced in this method. It has the ability of running for a long time without failure. Future investigation is needed to incorporate machine learning by heuristic algorithm in CASSARAM. Access to the new location and threat detection and prevention are not provided by CASSARAM.

Fuzzy set-based sensor search: In sensor similarity searching using fuzzy sets, sensor gives exact outputs as earlier. The technique can be used for realistic application such as searching videos and images. Sensors give output as per user demands satisfying user's criteria. Suppose, if a user wants to know the places with same climate condition, some temperature sensors with the same output will be generated. This method does not support provision for sensing as a service, machine learning, and threat detection and prevention. This method is not highly scalable but efficient in this sense that it returns the search result to the user with a fast response. Fuzzy set represents output with a few bytes and needs more work to search among large number of sensors.

Bayesian network-based sensor search: This method supports user's criteria. Bayesian network is essential for forming perfect prediction of desired sensors. This method supports realistic applications. Access to the new location is supported here. Provision for sensing as a service is not supported here. Sensors are read out randomly until it reaches to a point where matches can be found for a given number. Remote sensor readouts can be reduced accurately in this approach, and it improves searching for sensors that exhibits a given state when it is time for query for forming a prediction model. High scalability, machine learning, and threat detection are not provided here.

Distributed regression: User can define his requirements in this approach. This method is useful for drawing out more absolute information regarding the shape and structure of sensor data which ensures accuracy. Network nodes can fail due to lossy wireless transmission. Luckily, the distributed regression can be formed robust to such failures. This algorithm is really scalable as it can deal with large sensor networks consisting of hundreds of sensors. By building contour plots, analyzing number of sophisticated in-network applications, it is possible to have practical implications. This method does not support provision of sensing-as-a-service and machine learning techniques.

Shodan search engine is used to ensure threat detection and prevention by introducing IP exposure notification system. The whole working mechanism can be carried out using real-life IoT devices so it supports realistic applications. The system can work with millions of IoT devices so it is definitely scalable. Access to the new location and machine learning technique is not supported in this system. The system prevents the threats for exposed IoT devices. Exposed IP address can be found, and the level of exposure can be estimated in this system. Thus, the risk can be reduced. IP exposure threats can be reduced in state institutions, ranging from personal to national.

5 Open Research Area

Though the existing techniques can provide us with the best utilization of the sensor networks deployed in IoT, the scope of future research can be elaborated to the following sectors:

Scalability: Scalability in wireless sensor network (WSN) actually refers to scaling a network with a range of high node numbers and a bit of high node density [9, 10]. Implementation with global sensor networking (GSN) system may be approachable for developing a generic platform for deploying sensor networks and processing data produced by the sensor network in a distributed fashion.

Sensing as a service: Since humongous devices and sensors are getting connected to the Internet, IoT is becoming a major topic in this technology driven era. Some major vital measurements can be obtained through device management, the network, and the collected data. The management and provisioning of such sensor devices and data can create new business opportunities with some new challenges. Industry and academic—both the sectors—need to manage the devices that are interconnected and utilize the opportunity represented by the huge amount of data generated. Again, sensor network infrastructure involves huge investment as well as high maintenance cost. It holds users back from setting up their own IoT systems and Web applications that exploit sensor data. Mobile devices or sensor networks are occupied with sensor data through some sort of operations so that the data supplied by the sensor providers can be subscribed and published. Cloud-based sensing as a service (CSaaS) can be proposed as it can engage and manage various types of sensors on IoT devices by using the layer of virtualization; for in-depth analysis regarding cloud system, data is collected from sensors and sensor networks. In SaaS approach, real time and historical data can be accessed for analysis, and it can be referred as a convenient and standardized sensing data service for the consumers and third-party applications that are data-dependent. Dynamic provisioning for users is enabled here to leverage the huge pool of resources on demand [11].

Machine learning: Machine learning is becoming a major key point in IoT enhancing artificial intelligence between IoT devices as well as resulting in energy savings. WSN consumes a major percentage of its energy. Specially, for higher required accuracy, higher computational requirements are needed, hence higher energy consumptions occur [12]. For large-scale energy-constrained sensor networks, transmitting all data directly to the sink cannot be considered as an efficient process. There is an effective and efficient solution for passing data to a local aggregator which is termed as a cluster head (CH) [13]. CH collects data from all the sensors within its cluster and then transmits data to the sink. Proper CH selection is very significant as it will reduce energy consumption and also enhance the network's lifetime [14]. Distributed machine learning techniques may be suitable for limited resource devices such as WSNs. Distributed machine learning methods need less computational power and also smaller memory footprint, compared to centralized learning algorithms [15]. This technique can help selecting proper CH, and

through this method, the system may be energy efficient and machine learning may be enhanced.

Threat detection and prevention: It is crucial to improve the process of threat detection and prevention in IoT as a numerous number of sensors are connected. However, it is becoming challenging task because there are many factors that can cause faults and can influence data.

Body area network: Some intelligent sensor nodes should capture various physiological signals such as—sensing, sampling, processing, and communicating physiological signals. The sensor used in wireless body area network (WBAN) would have to be low energy consuming, less complex, having smaller form factor as well as light in weight, power efficient, user-friendly, and reconfigurable. Data existing on multiple mobile devices and wireless patient nodes needs to be collected and analyzed in a seamless fashion. ZigBee (802.15.4) or Bluetooth (802.15.1) can be used for designing the nodes to sample vital signs and transferring relevant data to a personal server through a wireless personal network. Different platforms with variety of wide area network (WAN) access possibilities for Internet access can be used to accommodate a personal server application. [16]. Selection of platform is system-specific and needs to be selected to reduce obtrusiveness for a given user. An algorithm can be proposed for routing in WBAN.

6 Conclusion

In the near future, billions of objects are going to be included in IoT. A lot of sensors are expected to be attached with those objects resulting in increasing number of sensors around us. To get desired data from those sensors and processing them for future applications, knowledge of appropriate sensor searching method is very crucial. For sensor searching criteria, sensors should be paired with IoT applications. Clustered sensor networks raise the challenge of searching and selecting sensors appropriately. For any query, the algorithm should search and select the right sensors for capturing data and combining together with numeric and semantic reasoning. The algorithm in a searching technique should be defined the in such a way so that the users do not need to access all the sensors for collecting data. Again, search algorithms can broaden the research area by focusing on the priority issues. Categorization of searching results on the basis of priorities can be added as a major feature of the search engines. From this research, the users can be benefitted by getting a clear concept of the overall analysis and comparison of the existing techniques including sensor search, sensor similarity search, and providing security.

References

1. Ashton, K. (2009). That internet of things thing. *RFiD Journal, 22*(7), 97–114.
2. Ko, Y.-S., Ra, I.-K., & Kim, C.-S. (2015). A study on IP exposure notification system for IoT devices using IP search engine shodan. *International Journal of Multimedia and Ubiquitous Engineering, 10*(12), 61–66.
3. Farin, N. J., Rahman, A., Mansoor, N., & Hossain, S. (2016). Wotcoms: A novel cross-layered web-of-things based framework for course management system. In *Proceedings of the First International Conference on Advanced Information and Communication Technology (ICAICT-16).*
4. Zhou, Y., De, S., Wang, W., & Moessner, K. (2016). Search techniques for the web of things: A taxonomy and survey. *Sensors, 16*(5), 600.
5. Perera, C., Zaslavsky, A., Christen, P., Compton, M., & Georgakopoulos, D. (2013). Context-aware sensor search, selection and ranking model for internet of things middleware. In *2013 IEEE 14th International Conference* on *Mobile Data Management (MDM)* (Vol. 1, pp. 314–322). IEEE.
6. Truong, C., Romer, K., & Chen K. (2012). Fuzzy-based sensor search in the web of things. In *2012 3rd International Conference on Internet of Things (IOT)* (pp. 127–134). IEEE
7. Mietz, R., & Römer, K. (2011). Exploiting correlations for efficient content-based sensor search. In *SENSORS, 2011*. IEEE.
8. Guestrin, C., Bodik, P., Thibaux, R. Paskin, M., & Madden, S. (2004) Distributed regression: An efficient framework for modeling sensor network data. In *Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium* (pp. 1–10). IEEE.
9. Liao, W.-H., Chen, C. C. (2010). A multi-dimensional data storage algorithm in wireless sensor networks. In *Network Operations and Management Symposium (NOMS).* IEEE.
10. Rumín, A. C., Pascual, M. U., Ortega, R. R., & López, D. L. (2010). Data centric storage technologies: Analysis and enhancement. *Sensors (Basel),* 10(4): 3023–3056. Published online 2010 Mar 30. https://doi.org/10.3390/s100403023.
11. Kim, M., Asthana, M., Bhargava, S., Iyyer, K. K., Tangadpalliwar R., et al. (2016). Developing an on-demand cloud-based sensing-as-a-service system for internet of things. *Journal of Computer Networks and Communications*, 3292783, 17.
12. Meshram, N. A., & Thakare, V. M. (2015). Secured wireless sensors network using machine learning approach. In *IEEE Conference Paper, Advanced Technologies in Computing and Networking*.
13. Tasmim, S., Kamal, A. H., Tusher, M. O., & Mansoor, N. (2020). DEB: A delay and energy-based routing protocol for cognitive radio ad hoc networks. Algorithms for intelligent systems. In *Processing. of International Joint Conference on Computational Intelligence* (pp. 643–654). Springer.
14. Mansoor, N., Islam, A. M., Zareei, M., Baharun, S., & Komaki, S. (2016, November). A novel on-demand routing protocol for cluster-based cognitive radio ad-hoc network. In *2016 IEEE Region 10 Conference (TENCON)* (pp. 632–636). IEEE.
15. Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H.-P. (2014). Machine learning in wireless sensor networks: Algorithms, strategies, and applications. *IEEE Communications Surveys & Tutorials,* 16(4) 1996–2018. arxiv.org/abs/1405.4463
16. Otto, C. (2006). An Implementation of a wireless body area network for ambulatory health monitoring. Huntsville, Alabama

SIGSv2: Enhanced Student ID Generator and Management System

Ian Benitez and Philip Alger Serrano

Abstract Electronic ID systems in educational institutions had advanced technologically and are used for integration to other information systems. To utilize this advantage, an automated generation and management of ID cards was developed. In its implementation, however, it was found out that the system lacked capabilities needed for a more effective process. So, an enhancement had carried out to improve the system's capabilities. Data was gathered using interviews, process observations, and questionnaire. The respondents evaluated the acceptability of both old and the new systems in terms of efficiency, reliability, functionality, and usability based on ISO 9126 software metrics, and they agreed that the new system exhibited superior performance. Further, comparison results revealed that the new system significantly improved in terms of efficiency, functionality, reliability, but had not achieved significant improvement in usability, particularly, on the user interface. With the deployment of the enhanced system, a more robust solution in generating and managing identification card was achieved that led to a more efficient delivery of service.

Keywords Software enhancement · ID generator and management · Process automations · SIGSv2 · Software quality

1 Introduction

Process automation in the frontlines of many higher learning institutions aims at increasing its productivity and bringing better services to its clientele. Improving customer service and efficiency is at the top priority of these establishments, and hence, the application of multiple technologies and process innovations is very essential to support the needs and expectations of the clients [1]. Currently, people have

I. Benitez (✉) · P. A. Serrano
Camarines Sur Polytechnic Colleges, Nabua, Camarines Sur, Philippines
e-mail: ianbenitez@cspc.edu.ph

P. A. Serrano
e-mail: pamserrano19@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_75

come up with more innovative ways in which the identification card (ID) system can be implemented. ID is very essential to identify students and staff who belong to the school community. ID card systems also help schools to provide a secure learning environment through its seamless integration in the student's transactions. Additionally, electronic ID systems became a modern tool used by many individuals in the community to accomplish an electronic-based transaction [2].

Camarines Sur Polytechnic Colleges (CSPC), as a higher education institution, has been on an upward curve to adopt technologies to streamline some of their important operations. In 2014, an application called Student ID Generation and Management System (SIGSv1) [3] was developed and deployed in the College. The system generates electronic ID (e-ID) which is used by the students to link their information to a number of information systems such as the enrollment, student information system, and queuing system. However, implementation of the system revealed that it had many inefficiencies in its processes. Hence, there is a need to redesign the system to provide a more robust solution in generating and managing e-IDs. The objectives of this study are three-fold: firstly, to have an in-depth analysis of the performance of the old system, secondly, to develop an enhanced software application that will resolve the problems encountered in the old system, and lastly, to evaluate the level of acceptability of both systems applying ISO 9126 software quality metrics, and to determine their significant difference in performance.

2 Related Literature

ID system is an essential part of the services offered in any organization. The scheme started from simple ID cards, just become more sophisticated with the advent of technology. e-ID can be integrated with barcodes, RFID, and QR technology. To ensure the consistency and privacy of the study data in epidemiological studies, [4] proposed an ID generator that creates a unique identifier which not only carries a random identifier for a study participant but also supports the creation of structured IDs. Office automation like the ID system simplifies the tasks while improving productivity [5]. It is imperative, however, that the current system's limitations, which hinder its full automation potential, must be taken into consideration and must be innovated for more efficient performance. In the aspect of system development, the software development life cycle (SDLC) is a structure imposed in the development of a product [6]. Agile model is based on iterative and incremental development where requirements can be changed based on the customer needs [7], where adding functionalities to the previous release that produces software builds a very essential process.

3 Materials and Methods

This study was anchored on a descriptive-developmental research design. The proponents evaluated the old system to find out its existing problems pertaining to the processes involved in its operation and was enhanced anchored on the discovered problems. The two systems were evaluated in terms of ISO 9126 software quality metrics and were statistically compared to determine significant performance differences.

3.1 Software Development

Adopting Agile model in the system development, the researchers first looked into the detailed operations on each transaction and conducted observations and interviews with the users and students. Then, the collected data was analyzed. The researchers classified the problems and limitations encountered into several metrics namely: efficiency, functionality, reliability, and usability. Table 1 presents the problems of the old system.

In the coding phase, then proponents ensured a constant communication and feedback from the end users and iteratively build the solution based on their recommendations. The product was sent to the client for evaluation, feedback, or change request and was finally installed to the clients' computers.

Paradigm of the Enhanced System. The back-end database is managed by MySQL server, while the front-end was developed in visual studio. Other essential devices present are the e-signature pad and ID printer. An online form was integrated for easier collection of students' information. Figure 1 illustrates the framework of the system.

Table 1 Problems encountered in SIGSv1

Software quality metrics	Problems encountered
Efficiency	• Slow searching of records/information
Reliability	• Duplicate student records, periodic runtime errors, database anomalies, incomplete data validation
Functionality	• Absence of basic reports for the transaction such as the number of transactions per day, number of printed IDs, and tracking of claimants, no backup/restoration, background changer and back-to-back printing no available, manual students' profiling
Usability	• Some buttons and icons are not properly organized

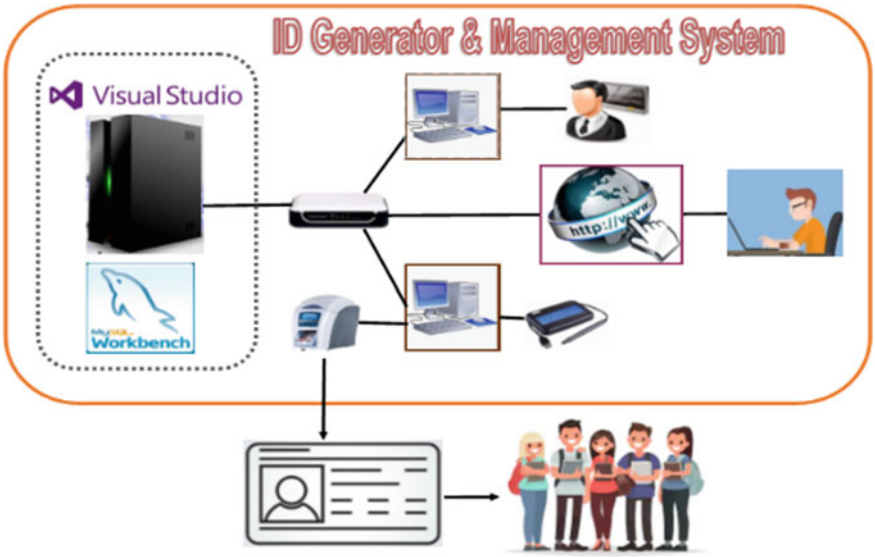


Fig. 1 System paradigm

3.2 Evaluation

The first set of respondents consisted of the incoming freshmen students from the different departments. The second set was composed of twenty-two (22) IT experts from the two colleges. The students assessed only the two (2) ISO 9126 metrics (functionality and usability) of the enhanced system, while the IT experts assessed the software quality of the old and the new system in terms of efficiency, reliability, functionality, and usability.

Sampling Method. Yamane's formula [8] was utilized in determining the sample size of the students. The confidence level was set to 95% and a 5% degree of error. The respondents were composed of the first-year enrollees of the College. Out of 2393 freshmen students, we determined 343 respondents.

Statistical Analysis. The collected data was analyzed using descriptive statistics in the form of frequency and percentage in the evaluation of the quality of the existing system and the new system.

Performance Comparison using a T-test. A t-test was used to determine if there is a significant improvement in the new system over the old system in terms of efficiency, functionality, reliability, and usability. This evaluation process was conducted utilizing the IT experts' responses only. The reasons why the students were not assigned to evaluate the old system were due to some constraints, particularly time, schedule, and resources. The student respondents were freshmen, and so, the best timing for software evaluation period was during their application for student ID. At that time, however, what they can utilize is the new system only through the

online profiling form. Providing an experimental setting for students to test the old system would be very difficult to attain due to lack of time and resources during the course of the study. Nevertheless, the same can be very easily accomplished by the IT experts, since they are the direct users of both systems. In light of this fact, the researchers agreed upon that the IT experts are going to evaluate both systems, while for the students, only the new system would be evaluated.

4 Results and Discussion

The solutions to problems encountered in the analysis phase are presented in Table 2.

4.1 The Enhanced System (SIGSv2)

Shown below are some screenshots of the developed front-end of the enhanced system.

The main screen of the system, shown in Fig. 2, was not a total revision of the old system. Although it exhibits a similar interface, a number of functionalities were added to the system. Figure 3 is the main form used to finalize ID printing. On the left pane is the student's information collected via online profiling or via local encoding by the application user.

Table 2 Crafted solutions based on the problems encountered

Software quality metrics	Solutions
Efficiency	• Enhancing the query statements, adding values as parameters.
Reliability	• Back-end database restructuring, get rid of illegal operations, improved data validation
Functionality	• Essential reporting and backup/restoration functions, online form, ID-background changer, back-to-back printing, self-profiling
Usability	• Improve buttons placement and feedback organization

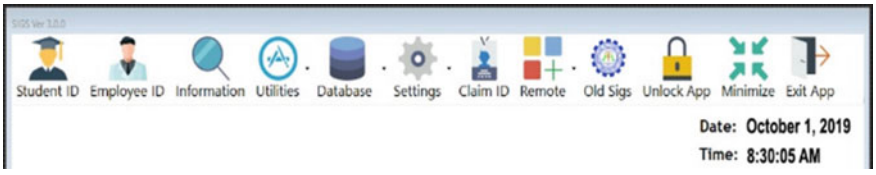


Fig. 2 Main screen

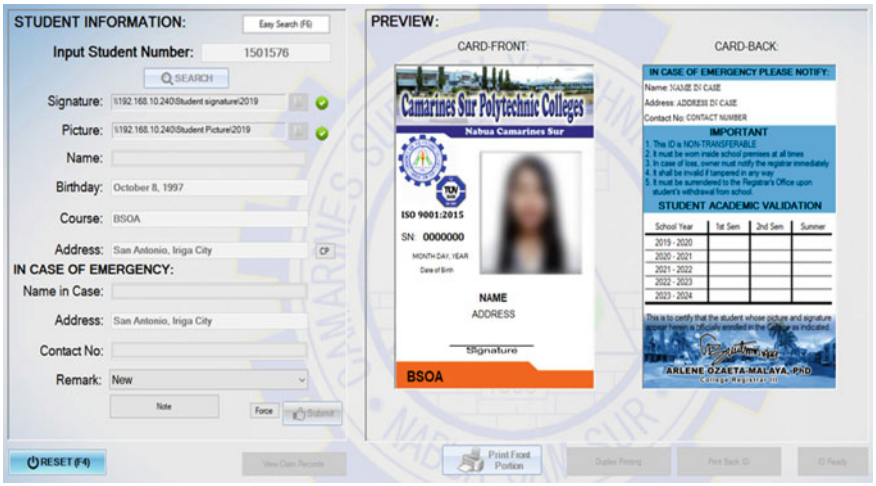


Fig. 3 ID generating form

4.2 System Evaluation Results

IT Experts Evaluation. Utilizing four ISO 9126 criteria for software quality, Fig. 4 shows the performance rating of the old and the new system in terms of efficiency, functionality, reliability, and usability. In the assessment of the software quality of the two systems in terms of efficiency, the old system achieved a weighted mean of 2.42. On the other hand, the respondents rated the new system with a weighted mean of 4.47. The notable difference between the two systems was about the response time to the user's query. As for functionality, respondents rated the old system with a weighted mean of 2.33, while the new system was rated with a weighted mean of 4.50. The reliability test result confirmed that the old system's processes are not stable enough to be a reliable software for ID generation and management. The respondents had an overall rating of only 2.34, which means they disagreed that the old system possesses these criteria for reliability. On the other hand, the reliability test results of the new system showed that the respondents agreed on all the given criteria for reliability, having an overall total rating of 4.23. For the usability rating, experts rated the old system 3.97, while the new system was rated 4.6. The overall usability rating particularly on the user's interface expressed that there is only a slight improvement in the new system.

Students' Evaluation. Figure 5 shows the rating of the students in terms of efficiency and usability of the new system.

In terms of efficiency, the overall rating of 4.31 implies that the new system is working efficiently. In terms of usability of the online profiling form, the new system passed the usability test, with a weighted mean of 4.36.

Significant Difference in the Experts' Evaluation of the Two Systems. It can be gleaned in Fig. 6 the difference in the performance of the two systems in terms

Fig. 4 IT experts' evaluation

of efficiency, functionality, reliability, and usability. The result of the comparative evaluation revealed that the new system had significantly improved over the old system in terms of efficiency, reliability, and functionality. On the usability criterion, however, it was determined that the performance difference between the two is very minimal and was deemed insignificant despite giving a higher rate on the new system by the IT experts.

4.3 System Deployment

Eventually, the finished product was deployed during the first semester enrollment, the school year 2019–2020, where the new system was installed in the MICT office as a frontline service. Users were trained, and the flow and stability of the software and its infrastructure were maintained. Figure 7 shows staff using the new system to process an ID card.

At the same time, the online profiling form was utilized by the students. Figure 8 shows students using the Web-based profiling form to fill up their basic information for the ID generation.

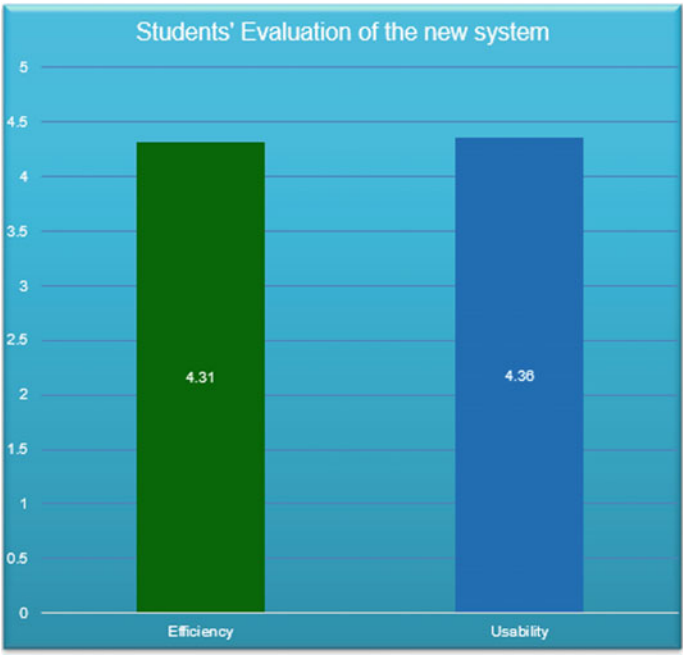


Fig. 5 Students' evaluation of SIGSv2

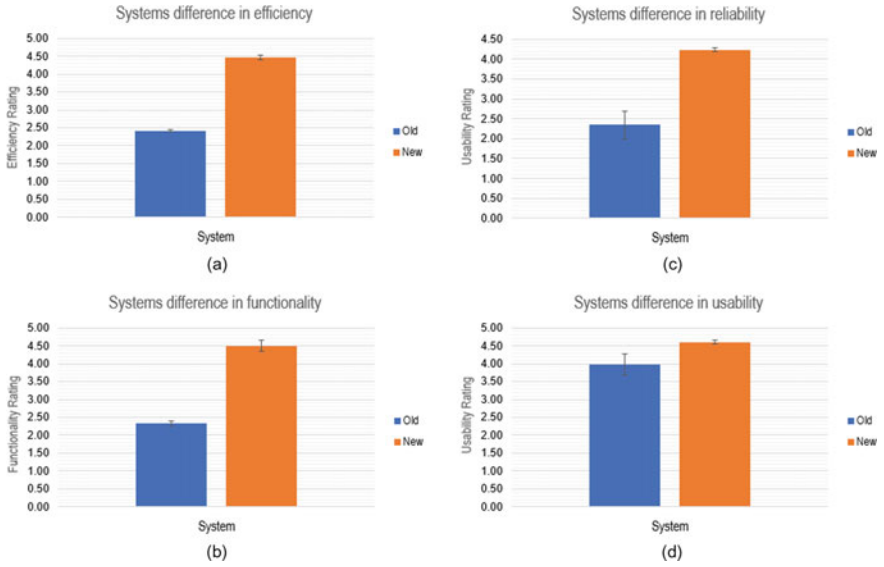


Fig. 6 Performance difference between SIGSv1 and SIGSv2

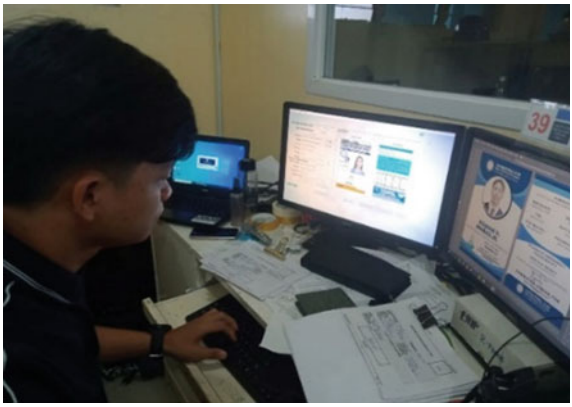


Fig. 7 Personnel processing the ID card

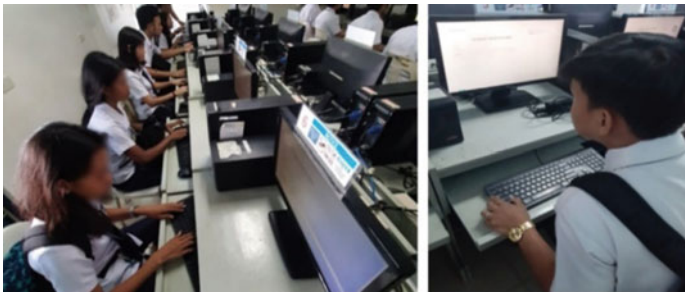


Fig. 8 Deployed system utilized by the students

Upon claiming the ID, the students must sign through a digital signature pad before ID card printing takes place. Figure 9 shows the sample signature of a student taken through the device.

Fig. 9 Taking signature with a signature pad

5 Conclusion and Recommendations

In this study, an ID generator and management system was enhanced to improve the process of generating and managing electronic IDs for students. An online student profiling form was integrated to the enhanced system to avoid long queue of students due to manual profiling process. To evaluate the software quality level of both the old and the new systems, they were both rated by the experts using the ISO 9126 software quality standards. Based on the results, the old system generally did not meet satisfactorily the criteria on efficiency, functionality, reliability, and usability. On the other hand, the new system met the good software quality standard when evaluated by the respondents. The comparison between the new and old systems performed by the experts revealed that there is a significant improvement in the new system over the old system in terms of efficiency, functionality, and reliability, while there is no significant difference in the two systems in terms of usability, particularly on the user interface. At the same time, the students' evaluation of the new system revealed that the new system passed the software quality standard in terms of efficiency and usability. The new system can be extended through a provision of additional system features such as back-to-back ID printing and improving resiliency in times of data loss or system errors.

References

1. Rains, J. (2014). Improving efficiency and customer service in higher education, no. AUGUST, (pp. 1–4).
2. Omorog, C. D. (2019). LIDS: Local ID system for e-government systems. *Indian Journal of Science and Technology, 12*(24), 1–7.
3. Nuñez, J. L. V., Ordoñez, A. J. C., Prestado, J. F., & Virtus, C. M. O. (2014). CSPC-students identification card generating system, no. October.
4. Olden, M., Holle, R., Heid, I. M., & Stark, K. (2016). IDGenerator: Unique identifier generator for epidemiologic or clinical studies. *BMC Medical Research Methodology, 16*(1), 1–10.
5. Bhuyar, R., & Ansari, S. (2016). Design and implementation of smart office automation system. *International Journal of Computers and Applications, 151*(3), 37–42.
6. Sarker, I. H., Faruque, F., Hossen, U., & Rahman, A. (2015). A survey of software development process models in software engineering. *International Journal of Software Engineering its Applied, 9*(11), 55–70.
7. Weintraub, J. (2006). Implementation of an Iterative and Incremental SDLC (Systems Development Life Cycle) Model Development Project for a Financial Services Organization.
8. Joskow, J., & Yamane, T. (1965). Statistics, an introductory analysis. *Journal of American Statistical Association, 60*(310), 678.

Motion Control and Sensor Fault Diagnostic Systems for Autonomous Electric Vehicle

Raghavendra M. Shet, Nalini C. Iyer, P. C. Nissimagoudar, Akash Kulkarni, J. Abhiram, and S. K. Amarnath

Abstract Motion control of an autonomous or self-driving car depends intensely on data associated with its lateral and longitudinal motion sensors. These sensor data is used to formulate necessary and appropriate control signal for smooth and stable maneuver, through steering and acceleration subsystems. The work here focus on implementation of such motion control systems on a physical vehicle that is modified to achieve desired autonomous functionality on a predefined path. In addition, a sensor diagnostic system for monitoring the sensor health condition is tested to identify the faulty values in the system.

Keywords Steering · Acceleration · Sensor fault · Motion control

1 Introduction

While autonomous or self-driving cars are among the impending technology of twenty-first century, though being as old as the automobile industry. The technology behind this has been evolving for decades, tracing back to the automated highway system project, and before. The technology has matured that advanced driver assistance systems (ADAS) such as collision detection, cruise control, automatic lane keeping are part of standard fitment in number of vehicles. In addition, there are many different autonomous vehicle projects in various stages of design and development, including extended on-road testing of multi-vehicle fleets.

The autonomous-driving or self-driving system is getting progressively composite and must be proved effectively before deployment. Five major components of self-driving technology (i) Environment mapping which deals with creation of occupancy grid, localization map, and detailed map for the drivable area of the

R. M. Shet (✉) · N. C. Iyer · P. C. Nissimagoudar · A. Kulkarni · J. Abhiram · S. K. Amarnath
KLE Technological University, Hubli, Karnataka, India
e-mail: raghu@kletech.ac.in

N. C. Iyer
e-mail: nalinic@kletech.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_76

autonomous vehicle. (ii) Environment perception, which deals with localizing the ego vehicle along with providing visual perception. (iii) Motion planning, which involves creating global plan, behavioral plan, and local plan for the vehicle. (iv) Controller, which deals with longitudinal and lateral actuation in the form of throttle and steering control. Self-driving technology has a vast number of domains under it and covering the entire span in one research is not a feasible method; hence, choosing a domain which is important, has a lot of scope for development and makes an impact to the society becomes crucial.

Verification and validation of sensor data along with motion planning is a domain, which needs a lot of research and improvement as it deals directly with the sensor's hardware as well as software integrity. While processing progressively its effectively status has to be tested before deployment. The pledge of autonomous driving system and computation becomes the major part of the self-driving vehicle; keeping track of sensor working plays a very important role to provide transparency between the person and vehicle.

Hence, we will discuss the work contributing toward the integration of motion control and diagnostic systems over physical car for demonstration of autonomous functionality with appropriate mechanical actuations. Our work addresses the self-driving components related to motion planning, specific to lateral and longitudinal maneuvers/movement and associated controllers, that is actuation of steering and acceleration(motor) control along with monitoring and providing diagnostics for related hardware and software modules. The path planning algorithms are used for a predefined path using distance approach method and is used as reference.

Organization of rest of the paper is as follows. Section 2 deals with design and methodology of the proposed system, Sect. 3 deals with the results of motion planning mechanism and passive diagnostics of sensor. Conclusion is discussed in Sect. 4.

2 Design and Methodology

This section deals with planning lateral and longitudinal movements of ego car under given speed, localization map, and other conditions followed by actuation control with necessary interfaces along with providing diagnostics for related sensors and actuators.

Figure 1 is a functional block diagram of verification and validation of data, which explain how sensors interface with the main program and how the data is displayed over the graphical user interface.

Below is a functional block diagram of motion planning which includes the control of acceleration and steering, feeding back the data to the system to simulate the collected data.

There are two specific domains of this project: one focuses on steering and acceleration control as a part of motion planning and the second focuses on verification and validation of sensor data.

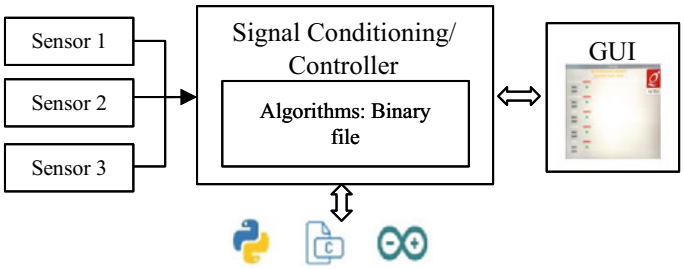


Fig. 1 Functional block diagram for verification and validation of data

2.1 Motion Planning

Motion planning is the term used in self-driving cars to find a sequence of valid configurations that moves the car/robot from the source to destination. A basic motion-planning problem is to compute a continuous path that connects a start configuration S and a goal configuration G while avoiding collision with known obstacles.

The block diagram representation motion-planning module with steering and acceleration control is shown in Fig. 2.

The proposed motion-planning module for the predefined path has two modules,

- Steering control and interfacing of pre-planned data (Lateral movement planning with steering control and interfaces)
- Accelerator control and interfacing of pre-planned data (Longitudinal movement planning with accelerator control and interfaces).

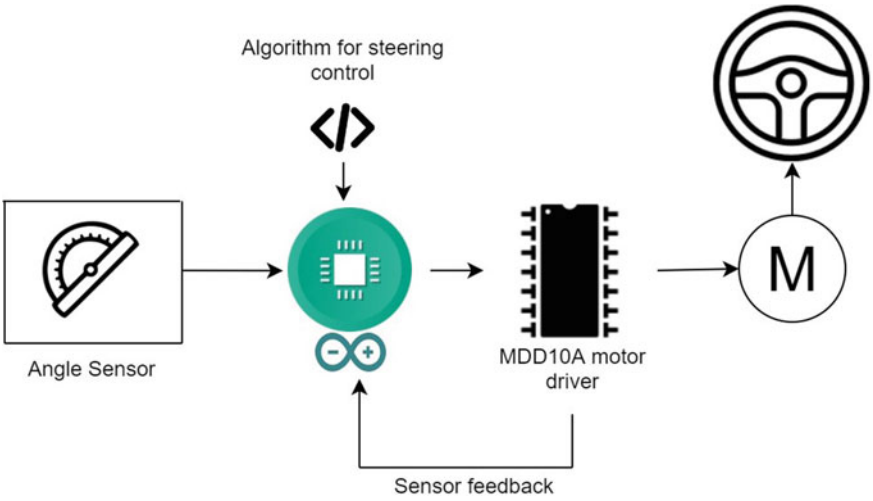


Fig. 2 Functional block diagram for motion planning

2.2 Steering Control and Interfacing of Pre-planned Data

This module consists of planning and controlling steering actuation according to predefined path for the set speed. The mechanism is showed in Fig. 3. The ego vehicle is designed to traverse the predefined path having certain fixed distance, and which is to be traversed by the vehicle considering possible maneuvers. The parameters related to predefined path are distance, diameter of the road, wheel dimensions, steering angle, speed and slope profiles. The control input is the steering wheel angle which is potentiometer fixed to the steering axle. Controlling of the steering motor was achieved with the help of MDD10A, Arduino and the power supply for the motor of 24 v.

For the desired lateral maneuvers, steering angle is the required input, which is priory stored in the look up table. Based on steering angle input, the potentiometer connected to steering wheel adjusts and accordingly DC motor gets energized and rotates. The associated gears in turn move the steering axle. To overcome the error between desired vehicle turn and actual vehicle turn, steering angle sensor input is provided as feedback.

Workflow:

Predefined path data is feed to the controller, which is an angle sensor values. According to the predefined data, the motor starts to rotate left or right taking an angle sensor as feedback, which is read by the controller. In addition, the process repeats.

- MDD10A motor driver: MDD10A is the dual channel version of MD10C that is designed to drive two-brushed DC motors with high current up to 10A continuously. Just like MD10C, the MDD10A also supports locked anti-phase and sign magnitude PWM signal. It is also using full solid-state components which result in faster response time and eliminate the wear and tear of the mechanical relay.

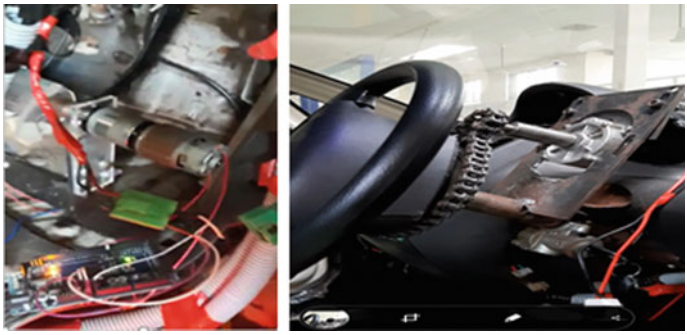


Fig. 3 Acceleration and steering mechanism

2.3 Accelerator Control and Interfacing of Pre-planned Data

This module deals with providing desired longitudinal control in terms of various speed profiles. The mechanism is showed in Fig. 3.

- The accelerator is nothing but speed control of the vehicle. Where we have connected the potentiometer to the accelerator input and fed the predefined data to it. The process is carried through the Arduino controller and potentiometer with 5 v of input.

2.4 Verification and Validation of Sensor Data

Verification and validation of sensor data are of very high importance when it comes to deploying sensors on an autonomous vehicle because there will be no human intervention and the vitals of sensors become the heartbeat of the vehicle. Malfunctioning of any sensors will result in a cascade of system failure that can result in huge consequences if not treated properly. Hence, validation and verification of data become extremely important in identifying and preventing the malfunctioning of sensory by throwing interrupt and completely immobilizing the vehicle.

- To prevent this from happening, diagnostic systems have to be implemented on the system that takes sensor data as input and analyses the sensor data for any faulty values or values which are different from other sets of normal values. Having done that, it classifies the data based on what kind of output is being obtained. If data is in the range of maximum and minimum values set by the manufacturer, then the system ignores it, else action will be taken.

2.5 Graphical User Interface

The graphical user interface is the face of all the verification and validation functionalities running in the background. It is developed in Python using the Tkinter library. Tkinter is Python's standard graphical user interface package. It is a thin object-oriented layer on top of Tcl/Tk. With the help of this library, it is able to create graphical figures and assign functions to it using class and objects.

3 Results

Vehicle is designed to traverse the predefined path. It is the path of certain distance that is to be traversed by the autonomous vehicle considering all the test cases.

Table 1 Parametres considered during implementations

S. No.	Parameters	Values
1.	Length of straight path (each side)	100 m
2.	Path angles	230, 240, 260, 335, 70, 48 (in degrees)
3.	Diameter of wheel	14 in.
4.	PWM values	0.06 (Low speed), 0.6 (Medium speed), 1 (High speed)
5.	Speed	0–10 km/hr

Parameters considered during implementation are represented in Table 1. The angle 230° is considered as the base angle, i.e., 0°.

Pre-planned path

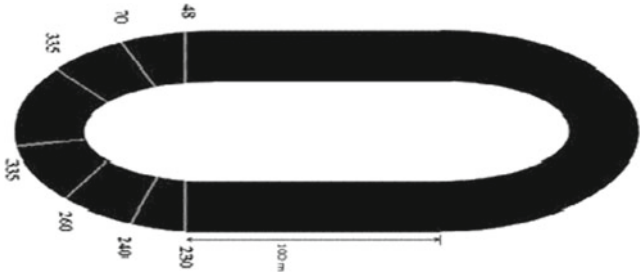


Figure 4 is the main page of the graphical user interface. We can see that there are several labels and buttons, which perform different tasks. The ignition status button triggers the ignition switch of the car, and when the ignition is on, the activity label displays whether the module is working properly or not with the help of algorithms explained above. Similarly, all the critical functionalities of the vehicle are associated with a fail-safe defense mechanism kind of system which keeps track of each and every functionality of the module and reports an error if any. This checks for any sensor faults used for desired motion control application by one-to-one monitoring of device.

Fig. 4 Graphical user interface

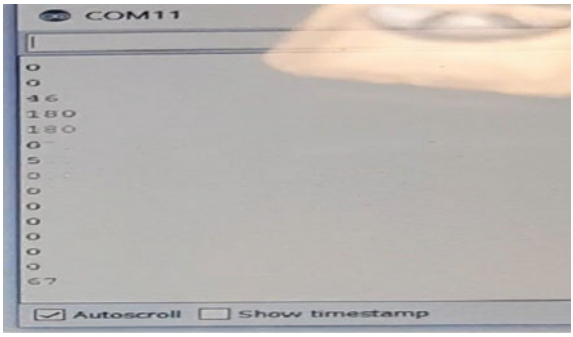


Fig. 5 Acceleration data simulation

A DC motor supported by a chain drive drives the of steering system. Figure 5 demonstrates the acceleration data simulation. The data of the wheel speed is recorded for a particular track and is stored in a file.

The collected data is then re-simulated over the course without the help of any human intervention and as the track remains the same, the vehicle mimics the speed of the previous collected model. This is a very important step to control the speed of vehicle during a curve or a speed-restricted area. The experimental result is illustrated in Fig. 6. The next step would be training the vehicle to record data on each movement and tag it with the particular geo location; hence, every time the vehicle passes the location, it will automatically find the speed restrictions set by the trained data.

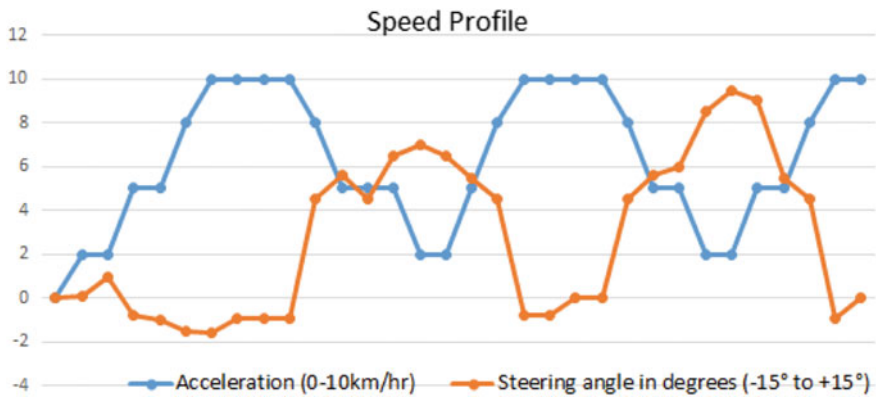


Fig. 6 Acceleration and steering sensor data

4 Conclusion

The vehicle traces the predefined path making it an autonomous vehicle using path planning algorithm predefined. The motion planning unit sends the steering angle in the curved path so that the vehicle takes the turn with the angle specified for steering. The braking unit applies the brakes on receiving the signal to apply the brake. Therefore, the vehicle behaves autonomous and traces the predefined path taking care of the various situations using the steering and acceleration modules. Also sensor faults are monitored using GUI. In the future, global positioning system (GPS) and inertial measurement unit (IMU) can be included in the algorithm so that it works not only for fixed paths but also for other paths.

References

1. Huang, W. H., Wang, K., Lv, Y., & Zhu, F. H. (2016). Autonomous vehicles testing methods review. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*.
2. Iyer, N. C., Gireesha, H. M., Shet, R. M., Nissimgoudar, P., & Mane, V. (2020). Autonomous driving platform: An initiative under institutional research project. *Procedia Computer Science, Elsevier, 172,* 875–880. https://doi.org/10.1016/j.procs.2020.05.126.
3. ISO 15765-2:2004 (2004). *Road vehicles—diagnostics on controller area networks (CAN) Part 2: Network layer services*. International Standards Organization.
4. Johanson, M., & Karlsson, L. (2006). A framework for distributed collaborative automotive testing. In *4th Workshop on Challenges in Collaborative Engineering*. Czech Republic: Prague.
5. Pooja, H., Priyanka, G., Prithvi, H., Sapna, U., Ramya, S., Priyanka, B., et al. (2016). Speed Control of DC Motor Using PID Controller. *International Journal of Industrial Electronics and Electrical Engineering, 4*(5), 5–7. ISSN: 2347-6982.
6. Larses, O. (2003) Modern automotive electronics from a dependable systems perspective. *Technical Report, Royal Institute of Technology*. ISSN 1400-1179.
7. Chikkamath, S., Vannal, N., Shet, R. M., Nissimgoudar, P. C., & Iyer, N. C. (2016). Experiential learning: learning through projects. In *2016 IEEE 4th International Conference on MOOCs, Innovation and Technology in Education (MITE)* (pp. 251–257). Madurai. https://doi.org/10.1109/mite.2016.057.
8. Huang, W. H. (2016). Autonomous vehicles testing methods review. In *Conference: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. https://doi.org/10.1109/itsc.2016.7795548.
9. O'Connor, P. D. T. (1988). *Reliability engineering*. Hemisphere Publishing Corp. ISBN 0-89116-684-X.
10. Shet, R. M., Iyer, N. C., Nissimgoudar, P.C., & Ajit, S. (2015). Integrated experience: Through project-based learning. In *Proceedings of the International Conference on Transformations in Engineering Education*. New Delhi: Springer.
11. Menzel, T., Bagschik, G., & Maurer, M. (2018). Scenarios for development, test and validation of automated vehicles. In *2018 IEEE Intelligent Vehicles Symposium (IV)*.
12. Iyer, N. C., Pillai, P., Bhagyashree, K., Mane, V., Shet, R. M., Nissimagoudar, P. C., Krishna, G., & Nakul, V. R. (2020). Millimeter-wave AWR1642 RADAR for obstacle detection: Autonomous vehicles. In *Innovations in Electronics and Communication Engineering. Lecture Notes in Networks and Systems* (Vol. 107). Singapore: Springer.

Virtual Simulation and Testing Platform for Self-Driving Cars

Nalini C. Iyer, R. M. Shet, P. C. Nissimagoudar, H. M. Gireesha, Venkatesh Mane, Akash Kulkarni, Ajit Bijapur, A. Akshata, and P. Neha

Abstract This paper is a document that delineates the autonomous functionalities on a simulation platform which helps a self-driving car by using a combination of different sensors like camera lidar, radar, and artificial intelligence using this sensor module that has to reach its destination without any human conduction and it has some functionalities like recognition, judgment, and operation offline testing remains a very important method permitting affordable and economical validation of car performance and vehicle management algorithms in multiple virtual eventualities hence to achieve the autonomous functionality.We propose the use of Autoware virtual simulation platform using robot operating system (ROS) which will provide the necessary modules like 3D map generation, localization, detection, prediction, planning, and control where we can generate a vector map which is necessary for path planning and object detection is achieved using the existing algorithms embedded into the project along with localization of the autonomous vehicle is also achieved in our research using NDT mapping by the means of it the position of the car, perception, and maneuver with the less error rate by using programming tools, and multiple test cases are generated through which validation of the functionalities is achieved on the virtual platform.

Keywords Autoware · Localization · ROS · Vector map · Object detection

1 Introduction

Autonomous vehicles and electric vehicles are the next generations of mobility infrastructure in the automotive industry. The industrialists and other companies are working on achieving the SAE levels of autonomy. The SAE stands for 'Society of Automotive Engineers' which defines six levels of driving automation. Level 0

N. C. Iyer · R. M. Shet · P. C. Nissimagoudar (✉) · H. M. Gireesha · V. Mane · A. Kulkarni · A. Bijapur · A. Akshata · P. Neha
School of Electronics and Communication Engineering, KLE Technological University, Hubli, Karnataka, India
e-mail: pcnissimagoudar@kletech.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_77

with no automation to level 5 with complete automation. Autonomous driving which is variably known as self-driving vehicle or robotic car is one of the most emerging technologies that has gained interest from industrialists and researchers because of its good deeds to the society and environment as well. These cars are a major advantage which reduces the accident rates, labor-cost savings, and reduction in carbon dioxide emissions through optimized driving. But the autonomous vehicles are restricted to operate on roads because of the safety regulations which are held by the government. Automotive Safety Integrity Level is a scheme defined by the ISO 26262 for the functional safety for road vehicles standard that needs to be passed by the automotive companies [1].

To build an autonomous vehicle, it is achieved in four main processes of perception, localization, control, and planning. In brief, perception is the first stage of the computational pipeline where a car has to detect vehicles ahead with its sensors and continuously track their movements. Localization, when self-driving cars use maps to figure out what the world is supposed to look like but it is not familiar with the car to understand. So to minimize that error, localization is done. The car has to later plan a path in finding a safe, comfortable, and efficient path to maneuver around the traffic. All of the above can also be achieved on simulation platforms and programming tools [2]. We can develop hands-on experience working with software programs. Simulations provide opportunities to engage developers, as they get to learn in an interactive environment that is close to the real-life situations. The computation also requires high-end GPUs and CPUs for faster processing and high power. There are many open-source platforms for self-driving and they are Apollo, Airsim, Neurojs, Carla, and Autoware. Autoware is a software that allows you to virtually simulate and show how an autonomous car works virtually. So, to plan motion and maneuver the vehicle, we need to generate an ADAS map and guide the vehicle through waypoints. To generate this ADAS Map, it takes software to tackle the problem of map generation for Autoware. The vector map states information of road structure and path planner uses a vector map to create the path to guide the vehicle [3] (Fig. 1).

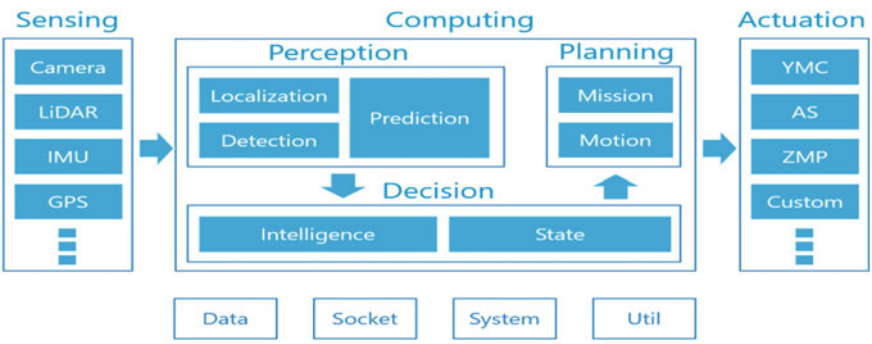


Fig. 1 Autoware architectural diagram

The above diagram depicts the architectural diagram of Autoware. Autoware is all in one open-source software for autonomous vehicles that provides modules for our dataset. The sensors on the left-hand side give input data after sensing. During the computation, it undergoes perception, planning, and decision making. These computations and decision making control the actuators like steering, throttle, etc. During computation, it has to localize to know the location of the vehicle. And also it has to detect the objects to take decisions accordingly [4].

2 Methodology

The implementation involves both software and hardware requirements set up to perform the task. The implementation was achieved on the Ubuntu 18.04 v Platform which also requires libraries and other dependencies that need to be installed further. As Autoware is based on ROS and we need to install Robot operating system (ROS) as well. So, on the basis of compatibility and requirements, the installation varies. Ubuntu 18.04 v requires ROS melodic as it is not compatible with ROS kinetic. To install ROS, we need to configure the Ubuntu repositories and setup the system to accept the packages from ROS.

The robot operating system (ROS) is designed to be a loosely coupled system in which a process is called a node, and one task is the responsibility of each node. Nodes communicate with each other by exchanging messages through the logical channels that are called topics. By using this publish/subscribe model, each node may send or receive data from the other node (Fig. 2).

Nvidia-Docker builds a bridge between GPU and repository. Basically, this repository has all the libraries and packages. So, Nvidia-Docker acts as a wrapper around the docker command line that funds a repository with the necessary dependencies to execute code on the GPU. As developing a docker speeds up the application that is why it is necessary. There are basically two versions: Docker CE (Community Edition) and Docker EE (Enterprise Edition). For small application computation, we need community edition. To install docker, we need to initially update the software

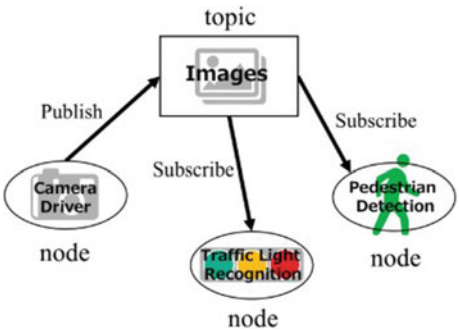


Fig. 2 Publish/subscribe model in ROS

repositories and uninstall the older version of docker. Later, install docker.io which is a docker image. Later, we also need Cuda, which is a computational platform and model developed Nvidia for computing on its own GPU. It speeds up the computation of intensive application. Cuda toolkits give a development environment, which includes libraries, debugging and optimization tools, compiler, and a runtime library to deploy the application. At last, we need to install Autoware which is built on ROS for the research and development purpose for autonomous driving technology. It is all in one open-source platform where you can achieve localization, detection, prediction, and planning through various algorithms with respect to those.

3 Implementation Details

3.1 Virtual Simulator in Autoware Environment

So, the applications achieved with these softwares are Visualization of Point Cloud Data(PCD) files from lidar, Visualization of ROSBAG from lidar, Localization through NDT Mapping using inbuilt GPS and IMU from lidar, interacting the lidar point cloud data with the camera data. Object detection is also achieved on this platform with other software dependencies like darknet and Yolov3 with Nvidia GPU and CUDA installed first visualize the ROSBAG and Point Cloud Data from Velodyne lidar on Autoware which we are able to see 3D map of the college. To plan the motion and to maneuver the vehicle, we need to create a predefined path to guide the vehicle so that it understands the shape and structure of the road. To generate such a map, we need a tool. The tool available as an open-source is Autoware's tools.tierIV that can create an ADAS map and waypoints to maneuver the vehicle. This tool requires point cloud data from the lidar using which we can view the 3D map of the college as well. This tool can define road shape, road surface, roadside, waypoints, road structure. In these, we can individually draw fences, road curves, edges, street light, footpath, crosswalk, lanes, and stop line as well. Vector map builders are quite difficult as they do not generate the map automatically. Basically, this is a sort of stitching from one node to another node, where each node has its own coordinates that x, y, z. We need to export the ADAS map, which produces the.csv file. This file has the coordinates values of each node. These files are required to generate the launch file for Autoware to view the map on Autoware. The only purpose is to generate a vector map to guide the vehicle [5] (Fig. 3).

The first application on this platform, localization which is achieved using a normal distribution transform (NDT) algorithm. The position of the vehicle can be located by scan matching with the 3D map of PCD files. Here, we break the point cloud map into 3D boxes called voxels. NDT follows three steps, i.e., Initialization, matching step, transformation determination. At the first stage, the environment of the PCD is divided into voxels which are also known as grids. It also initializes the grid size. The second stage generates a Gaussian distribution function of each voxel.

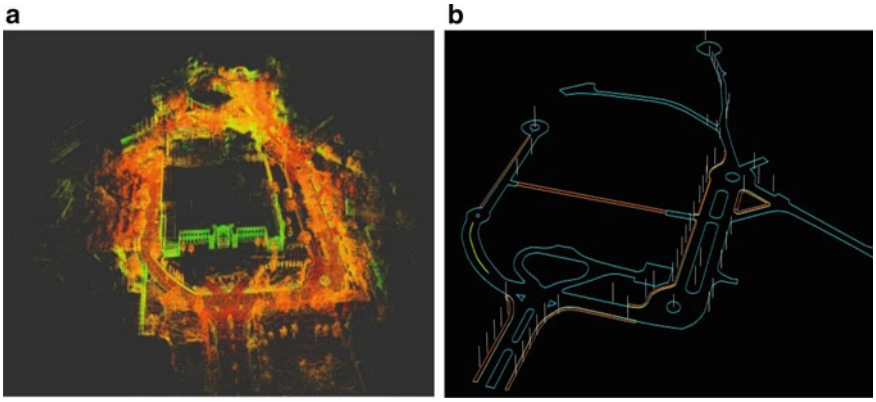


Fig. 3 **a** Dview of college map in PCD. **b** Vector mapping

For each point in the model, we search the voxel that verifies the distribution. In the third stage, to determine the transformation, we need to define the objective function. From this, we obtain *x*, *y*, *z*, yaw, pitch, roll values, and transformation matrix. The position error is around 10 cm.

Object detection, using an algorithm named YOLO 3 version. This is the updated version with an increase in accuracy and its speed. It uses a darknet neural network framework for training the images. This model processes images in real time at a speed of 45 frames per second. This algorithm is a concept of the bounding box which helps in the prediction of the object. Yolo predicts the bounding box using dimensional clusters in which the network predicts 4 coordinates for each box that is *Tx*, *Ty*, *Tw*, *Tz* with width and height of Pw and Ph, and if the cell is offset from the top left corner of the image by *Cx*, *Cy* The mathematical representation of the prediction is obtained using following equations.

The above image depicts the process of object detection. At very first, it divides the image into grids, i.e., $S \times S$ grid. Each grid looks for bounding boxes '*B*' along with the class probability '*C*' The whole prediction can be expressed in a single equation [6] (Fig. 4).

$$S \times S \times (B * 5 + C) \tag{1}$$

3.2 Connecting Autoware with LGSVL Simulator

Using this software, we have created a basic scenario like environmental conditions, for example, rain, fog, summer and also we can create pedestrians, traffic light control, collision test by using Python scripts and we can generate and validate it and this also provide real-time outputs from sensors including camera, LIDAR, RADAR,

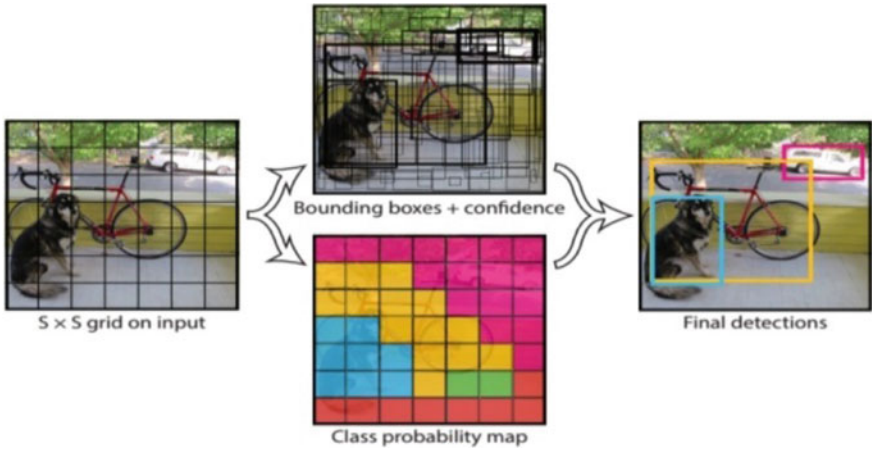


Fig. 4 YOLO detector results

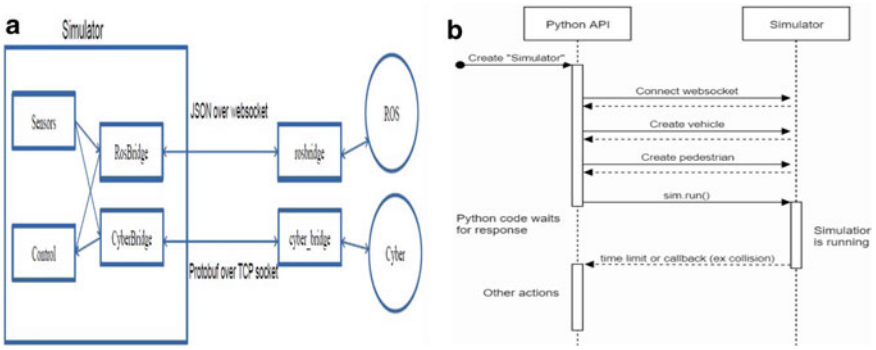


Fig. 5 **a** Working flow of ROS bridge connection. **b** TCP connection

GPS, and IMU [7]. Autoware communicates with the simulator using the rosbridge-suite, which provides JSON interfacing with ROS publishers/subscribers. The official Autoware docker containers have rosbridge-suite included (Fig. 5).

3.3 Experimental Setup

Experiments were conducted using Autoware V.1.12.0, CUDA 10, ROS Melodic, Nvidia Docker, Ubuntu 18.04. We used high speed computing GPU and it as specification is Nvidia GTX Geforce GPU(2080TI), Intel Core i7,32 GB RAM and also used logitech G29 steering wheel.

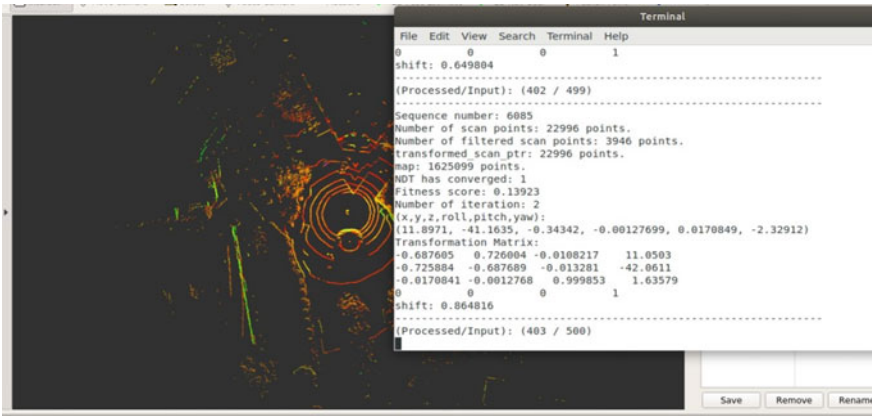


Fig. 6 NDT mapping in Autoware

4 Results and Discussion

This section elaborates on results of localization and object detection in autoware and results of LGSVL simulator.

4.1 Results of Localization in Autoware

The position of the vehicle can be found using the NDT algorithm. As we can see the output of this algorithm for localization using the NDT algorithm. This algorithm gives us the coordinates in the form of (x, y, z, roll, pitch, yaw). It also gives the transformation matrix. This position error of this localization is in and around 10 cm (Fig. 6).

4.2 Results of Object Detection

We can detect vehicles, pedestrians, traffic lights, persons, and a variety of objects in Autoware with maximum accuracy. The algorithm we have used is YOLOv3. In Fig. 6. The lines with different colors represent a different object. For example, the blue circles represent the vehicles whereas the green circles represent the pedestrians and yellow represents the curve edges (Fig. 7).

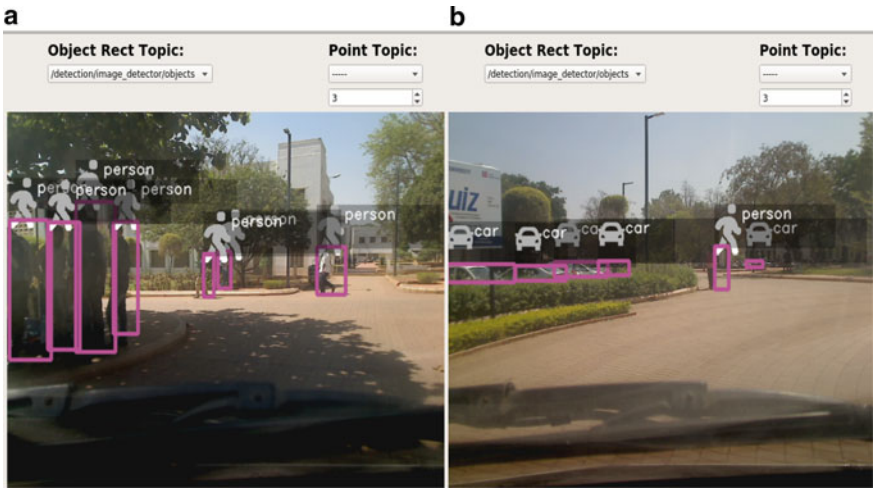


Fig. 7 **a** Person detection. **b** Car detection

4.3 Results of LGSVL Simulator

LGSVL Simulator exposes runtime functionality to a Python API which you can be use to manipulate object placement and vehicle movement in a loaded scene, retrieve sensor configuration and data, control weather, time state, and more (Fig. 8).

5 Future Scope

Some of the issues that can be further explored are

- In the future, we need to increase the computation speed and implement real-time service.
- The simulations can be extended further to explore and to achieve more of the parameters of Autoware.
- The simulations can be tested on many other simulation platforms as well.
- We can also use other development kits to deploy our simulations on to like Jetson AGX drivers.

6 Conclusion

Autonomous electric vehicles are the next generation of the automotive industry. We will get to see autonomous cars on roads as well. But, before real-time implementation, we need to test all the test cases on the simulation platform to see and analyze

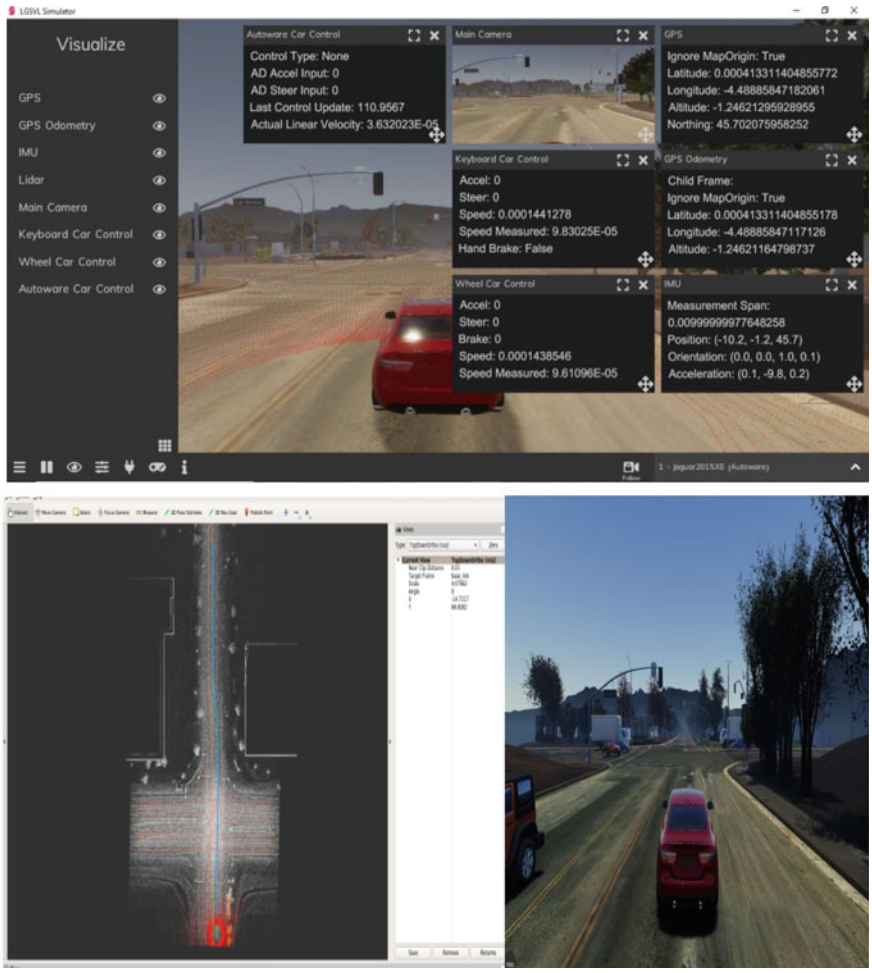


Fig. 8 ROS bridge connection between Autoware and LGSVL

how things work virtually than taking risks by bringing them on roads directly. The sensor's data like lidar (Velodyne), radar, and camera data is for sensing the environment and act accordingly. The lidar data is a major requirement in this project. The final simulations have met all the goals and objectives of the project. The localization has given the position coordination and in case of object detection, the detection is really fast with maximum accuracy. As well as the vector map is generated as expected. The ADAS map is bea. By the means of virtual simulation, we are able to successfully present the autonomous functionalities on the simulation platform.

References

1. Chavez, R., & Meystel, A. (1984). Structure of intelligence for an autonomous vehicle. In *IEEE International Conference on Robotics and Automation* (pp. 584–591).
2. Iyer, N. C., Gireesha, H. M., Shet, R. M., Nissimgoudar, P., & Mane, V. (2020). Autonomous driving platform: An initiative under institutional research project. *Procedia Computer Science, Elsevier, 172,* 875–880.
3. Sovani, S. (2017). Simulation accelerates development of autonomous driving. *ATZ worldwide Springer, 119,* 24–29.
4. Tun, W.N., Kim, S., Lee, J.-W., & Darweesh, H. (2019). Open-source tool of vector map for path planning in autoware autonomous driving software. In *IEEE International Conference on Big Data and Smart Computing (Big Comp)* (pp. 1–3).
5. Carballo, A., Wong, D., Ninomiya, Y., Kato, S., & Takeda, K. (2019). Training engineers in autonomous driving technologies using autoware. In *IEEE Intelligent Transportation Systems Conference (ITSC)* (pp. 3347–3354).
6. Chikkamath, Vannal, N., Shet, R. M., Nissimgoudar, P. C., & Iyer, N. C. (2016). Experiential learning: Learning through projects. In *2016 IEEE 4th International Conference on MOOCs, Innovation and Technology in Education (MITE)* (pp. 251–257.). Madurai.
7. Iyer, N. C., Pillai, P., Bhagyashree, K., Mane, V., Shet, R. M., Nissimagoudar, P. C., Krishna, G., & Nakul, V. R. (2020) Millimeter-wave AWR1642 RADAR for obstacle detection: Autonomous vehicles. In *Innovations in electronics and communication engineering. lecture notes in networks and systems* (Vol. 107). Singapore: Springer.
8. Chen, Y., Chen, S., Zhang, T., Zhang, S., & Zheng, N. (2018). Autonomous vehicle testing and validation platform. In *IEEE Intelligent Vehicles Symposium (IV)* (pp. 949–956).
9. Huang, R., Pedoeem, J., & Chen, C. (2018). YOLO-LITE: A real-time object detection algorithm optimized for non-GPU computers. In *IEEE International Conference on Big Data (Big Data)* (pp. 2503–2510).

Durable Implementation of Multi-cloud Storage with Assured Integrity for Sensitive Information

Shraddha Ghogare, Dheeraj Gupta, and Ambika Pawar

Abstract Cloud computing has gained an excellent response across several industries due to the agility and elasticity it offers. With the increase in the usage of cloud computing in both businesses and research industries, there was a significant ascent in threats and vulnerabilities. The cloud since, its inception, has a layered architecture with security deployed at each layer. Risks like access ambiguity, lack of central governance, data loss, data leakage, and data privacy are rising day by day. A multi-cloud strategy is to mitigate these risks and vulnerabilities. Even then, issues of data loss, privacy persist. This paper presents a methodology that reduces these limitations with the design and implementation of multi-cloud architecture and shuffling algorithm. The application developed provides security to the data deployed on the public cloud as well as maintains data reliability. This paper addresses the issues of data loss and theft. This design-pattern focuses on micro-services that also require reliability and availability of data.

Keywords Cloud computing · Reliability · Multi-cloud · Files · File upload · File download · Chunk shuffling · Durability · Confidentiality · Security

S. Ghogare (✉)
CSE Department, University of Texas at Arlington, Arlington, TX, USA
e-mail: shraddha.ghogare@mavs.uta.edu

D. Gupta
MIS, Eller College of Management, University of Arizona, Tucson, AZ, USA
e-mail: dheerajgupta@email.arizona.edu

A. Pawar
CSE Department, Symbiosis International University, Pune, India
e-mail: ambikap@sitpune.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_78

1 Introduction

There has been tremendous research done on the security of the cloud [10]. Initially, the focus was on providing role-based access control to users within the organization, and now it has also advanced to giving service-level accesses. Even then, few security incidences were reported [2, 9].

The system provides security and privacy to data stored on the public cloud. The project aims at securing cloud data by using a combination of shuffling algorithms, hashing algorithms, encoding/decoding techniques, and a multi-cloud storage approach. The users of the system are researchers, stakeholders, and industry professionals who require robust, secure, highly-available, and durable storage.

2 Related Work

An increasing number of enterprises are actively migrating their IT footprint from local data centers to the public cloud. However, the security aspects still need to be locked down in order to provide secure solutions to users. A new architecture type—a multi-cloud storage architecture that comprises of more than one public cloud provides enhanced security and integrity [6]. This approach first ingests in files to be uploaded on the cloud, and then it divides the files on the client system and then disperses the fragments on multiple public clouds as objects. This approach gained much popularity; however, data integrity and reliability were still a concern.

Reference [13] states that the users may dynamically enter and leave the cloud storage providers while the data would continue to be uploaded onto the cloud through background services. Redundant Array of Independent Disks (RAID) has been used for many years for various purposes because of its advantages like cost, safety, and enhanced performance. Such techniques can be utilized in a cloud-based environment, as well. Moreover, a multi-cloud environment can be realized by implementing this approach. A core limitation of the single-provider approach is the provider's policies, which the user has to agree upon unconditionally. This raises trust and acceptance issues among the users. This paper proposes a system that overcomes all these shortcomings effectually. The system has a central controller, which disperses the data to various cloud providers. This dispersion is done by using one of the algorithms which take care of the prevention of data loss, secret sharing, and management. Examples are: CORNFS, IDA [14], JigDFS [1], and Tahoe.

Reference [11] delineates an enhanced mechanism to provide privacy to sensitive data using file shuffling.

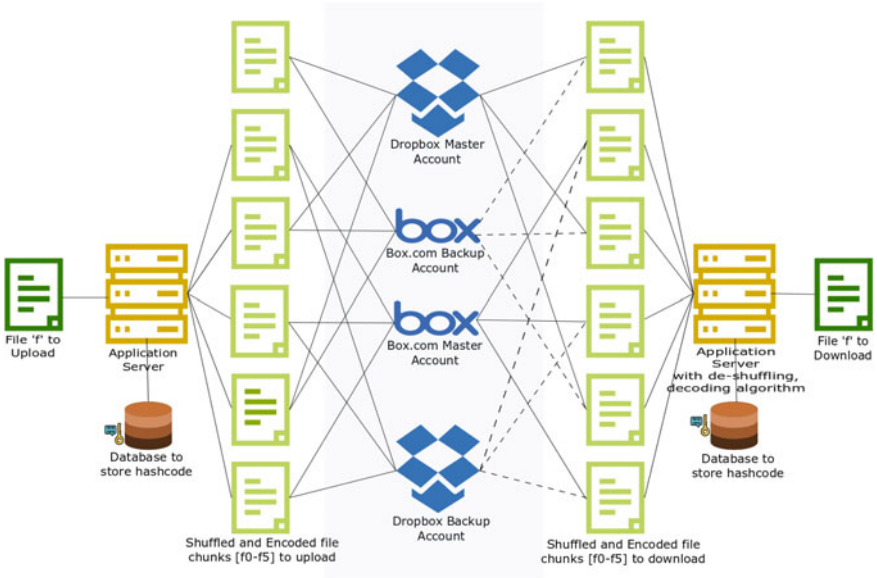


Fig. 1 Architecture of the system

3 System Overview

Considering the security requirements, this system tries to address and provide a one-stop solution. It primarily consists of a combination of multi-cloud architecture, a chunk shuffling algorithm, and encoding/decoding techniques. This design focuses on micro-services that also require the reliability and availability of data.

A hash generation algorithm ensures the integrity of the data. It generates a hash-code that servers as a key that can be associated with every file. The generation of hash-codes is done as the first step so that the hash code generated can be stored as a key. This key ensures the integrity of the file when it is downloaded by the users and regenerated from the shuffled and the fragmented chunks. A secure hash algorithm, SHA-256, is used in this project, which generates a 256-bit hash-code. The hash code has a message digest length of 256 bits. The data is processed in a total of 64 iterations.

The multi-cloud storage architecture is a scheme wherein more than one public cloud is used to store data to avoid downtime and security vulnerabilities. Data is deployed on several clouds and often is partitioned amongst multiple public clouds. The system does not replicate the entire file; instead, the file is split into the chunks, which are then replicated to assure the security and reliability of data on the public clouds. Furthermore, a shuffling algorithm adds an extra layer of security as it shuffles the chunks before they are uploaded to the public cloud. Files are first split into chunks, and these chunks are then shuffled. An individual can retrieve the file if they

have all the chunks of the file and the de-shuffling order. Hence, data confidentiality is assured. In addition to this, the shuffling algorithm is developed to work at byte level for all file making it file-type neutral. The details of the chunk shuffling algorithm are outlined in Sect. 4.

Data is vulnerable to attacks during transmission, which may corrupt the data. Therefore, the encoding of data in motion ensures that even if some of the data from file chunks is corrupted, the application server would be able to get back the original chunk of the file. Oracle and Apache both provide Base64 libraries to encode and decode string/byte data. The application and comparison of these two libraries are done based on the results obtained. This project presents a novel approach that uses multiple clouds, encoding, hashing, and shuffling algorithm to mitigate apprehensions related to data privacy. The viability, design, and implementation of these methods along with results obtained are described in this report.

4 Algorithms

As the number of chunks to de-shuffle is "n", the complexity of this algorithm is O(n).

Pseudo Code for splitting file into chunks

Input: 1. n: Number of chunks to be formed
2. fname: File in byte array
Output: Split chunks of a file
Algorithm:
Begin
1. Read number of chunks to be formed in variable n
2. Read file byte array to split in fname
3. Calculate chunk_size as chunk_size=file_size/n
4. Loop to create six chunks:
foreach *n* **do**
  Write right amount of bytes i.e. bytes of size chunk_size
  foreach *chunk_size* **do**
    Output to file_chunk number n
  end
end

Pseudo Code for chunk sequence shuffling algorithm

Input: 1. n: Number of chunks
2. f: File in bytes
Output: Shuffled array of sequence of chunks
Algorithm:
Begin
1. Read no of chunks, n
2. Read split and encoded file chunks in bytes
3. Shuffle chunk bytes as
foreach *n* **do**
  if $i\%2! = 0$ **then**
    odd_list = odd numbered chunks i.e. chunk number 1, 3, 5
  else
    even_list = even numbered chunks i .e chunk number 0,2,4
  end
end
6. final_list = odd_list.append(even_list)
7. Convert list final_list into a comma separated array to process further.

For splitting the file into 6 chunks, complexity is file size m. For shuffling the chunk bytes, two lists are created and appended, with complexity- number of chunks n. So, the total complexity is O(m+n).

Pseudo Code for chunk sequence de-shuffling algorithm

Input: 1. Returned list of file chunks from Download function
Output: File to be downloaded
Algorithm:
Begin
1. Create an array of bytes arranging the chunks in 0-5 sequential manner
2. Convert them to a list to pass to de-shuffle function
3. First three (index 0, 1, 2) are odd chunks and next (3, 4, 5) are even
4. Separate the lists in two - even and odd
foreach *no of chunks in list* **do**
  final_list=chunk_append(3, 0, 4, 1, 5, 2)
end
5. Convert final_list into array of bytes and file to download

Pseudo Code for uploading files from multi-cloud

Input: 1. s: Size of array
2. f: File chunks array in bytes
3. Dropbox master and backup account access token
4. box.com master and backup account access token

Output: Shuffled file chunks uploaded to multiple clouds

Algorithm:

Begin

1. Read shuffled array in m and initialize m = size of array;
2. Upload chunk bytes to the Dropbox directory as

foreach *m* **do**
  if $i\%2! = 0$ **then**
    Upload i to Dropbox's master account;
    if *success* **then**
      Goto upload to box.com Cloud Storage;
    else
      Upload i to box.com Cloud Storage's backup account;
    end
  else
    Upload i to box.com Cloud Storage's master account;
    if *success* **then**
      Goto step 5;
    else
      Upload i to Dropbox's backup account;
    end
  end
end

3. File chunks are uploaded without runtime errors to multiple clouds.

Pseudo Code for downloading files from multi-cloud

Input: 1. n: No of chunks to download
2. Dropbox master account with client API access key and secret key
3. Dropbox backup account with client API access key and secret key
4. box.com Cloud Storage's master account with developer token, client API key and secret key
5. box.com Cloud Storage's backup account with developer token, client API key and secret key
Output: list of chunks in an ordered sequence of odd-even chunks
Algorithm:
Begin
1.n=3
for *n* **do**
  2.Download files chunks 0,2,4 from Dropbox's master account
  if *success* **then**
    Go-to download from box.com Cloud Storage's backup account
  else
    Download chunks 0,2,4 from box.com Cloud Storage's backup account
  end
end
for *n* **do**
  3. Download files chunks 1,3,5 from box.com Cloud Storage's master account
  if *success* **then**
    Go-to step 4;
  else
    Download chunks 1,2,4 from box.com Cloud Storage's backup account
  end
end
4. De-shuffle the chunks as described in Algorithm 1

5 Performance Analysis

Figure 1 shows overall performance for file upload, and Fig. 2 shows the performance of file download operation of the system. Uploading speed is a little better than the downloading speed. However, the performance of the scheme depends upon the Internet speed user has. Today, as the data growth is enormous and files are of the scale of TBs and PBs size; the storage needed to store and backup these files is considerably high. With this scheme, it has been reduced drastically. For instance,

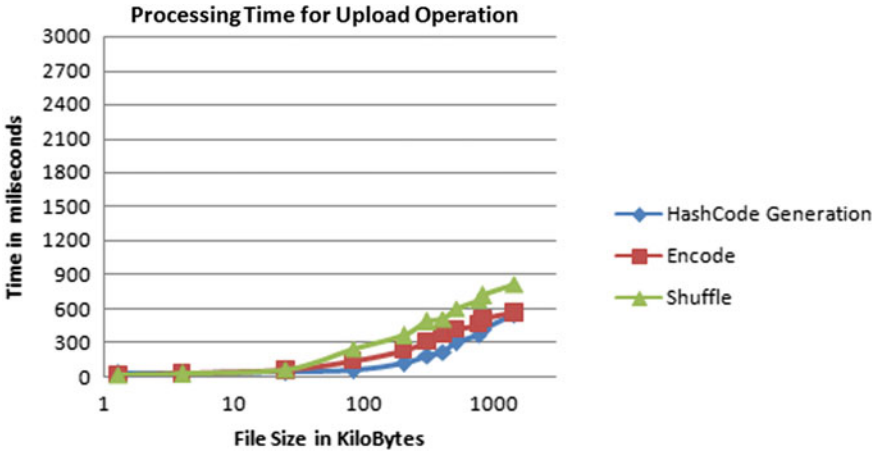


Fig. 2 Time taken by hashing, shuffling, and encoding algorithms

the storage needed to store and backup 2 MB files on a single cloud is 2 MB, and another 4–8 MB is required to replicate it on backup servers (considering file is replicated on 2 or 4 backup servers). The alternative is to use the scheme that makes effective use of multi-cloud architecture to improve the utilization of cloud storage.

5.1 File Upload Processing Time

See Fig. 2.

5.2 File Download Processing Time

The level of integrity, confidentiality, reliability, and availability is higher. An application server can be deployed on a private cloud to ensure scalability and effective load distribution. Additionally, the client-id and secret key generation for both Dropbox accounts and generation of developer token for box.com Cloud Storage accounts are being done manually due to the limitations of the public clouds. With the advancements in their APIs, the application server's IP can be configured to automate this, e.g., grant access to public clouds for requests coming from the application server for one year. A feature to upload multiple files done in a simple request can be implemented in the future (Fig. 3).

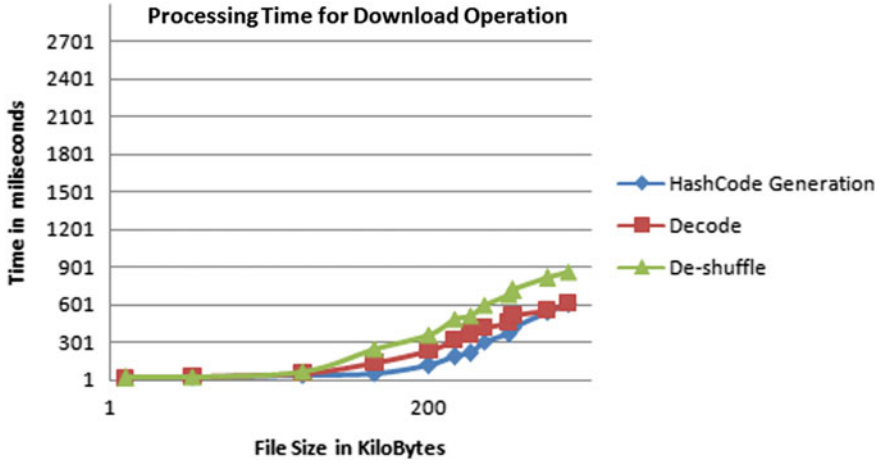


Fig. 3 Time taken by hashing, de-shuffling, and decoding algorithms

6 Conclusion

The system tackles the common problems in public cloud storage in terms of enhanced reliability, availability, confidentiality, and security of data. The system mitigates threats faced on the public cloud by micro-services by leveraging the multi-cloud architecture and a shuffling algorithm.

References

1. Bian, J. (2007). *JIGDFS: A secure distributed file system and its applications*. PhD thesis, University of Arkansas.
2. Cachin, C., Keidar, I., & Shraer, A. (2009). Trusting the cloud. *ACM Sigact News*, *40*(2), 81–86.
3. Chickowski, E. (2010). Flat-file databases often overlooked in security schemes. http://www.darkreading.com/risk/flat-file-databases-often-overlooked-in-security-schemes/d/d-id/1133363?.
4. Ghogare, S., Pawar, A., & Dani, A. (2015). Enriching utilization of cloud storage system. *International Journal of Applied Engineering Research*, *10*(10), 27413–27418.
5. Ghogare, S., Pawar, A., & Dani, A. (2016). Revamping optimal cloud storage system. In *Proceedings of 3rd International Conference on Advanced Computing, Networking and Informatics* (pp. 457–463). Springer.
6. Iacono, L. L., & Marnau, N. (2013). *Security and privacy-enhancing multicloud architectures*.
7. Vijayan, J. (2011). Vendors tap into cloud security concerns with new encryption tools. *Computerworld*.
8. Spillner, J. M. J., & Schill, A. (2013). Creating optimal cloud storage systems. In *Future generation computer systems* (pp. 1062–1072).
9. Kincaid, J. (2009, March). Google privacy blunder shares your docs without permission. *TechCrunch*.
10. Rhinosecurity Labs. (2019). Capital one cloud breach s3.

11. Pawar, A. V., & Dani, A. A. (2014). Design of privacy model for storing files on cloud storage. *Journal of Theoretical & Applied Information Technology, 67*(2).
12. Seiger, S. G. R., & Schil, A. (2011, September). Seccsie: A secure cloud storage integrator for enterprises. In *Commerce and enterprise computing (CEC)* (pp. 1062–1072).
13. Spillner, J., Bombach, G., Matthischke, S., Muller, J., Tzschichholz, R., & Schill, A. (2011). Information dispersion over redundant arrays of optimal cloud storage for desktop users. In *2011 Fourth IEEE International Conference on Utility and Cloud Computing (UCC)* (pp. 1–8). IEEE.
14. Zhang, X., & Wang, H. (2010). A study of the use of IDAs in cloud storage.

Hidden in Plain Sight: Building a Global Sustainable Development Data Catalogue

James Hodson and Andy Spezzatti

Abstract Modern scientific research for Sustainable Development depends on the availability of large amounts of relevant real-world data. However, there are currently no extensive global databases that associate existing data sets with the research domains they cover. We present the *SDG Data Catalogue*, an open, extensible, global database of data sets, metadata, and research networks built automatically by mining millions of published open access academic works. Our system leverages advances in Artificial Intelligence and Natural Language Processing Technologies to extract and organise deep knowledge of data sets available that is otherwise *hidden in plain sight* in the continuous stream of research generated by the scientific community.

Keywords Artificial intelligence · Open data · Information extraction · Entity linking · Natural language processing

1 Introduction

In 2015, the Unied Nations transitioned from the Millennium Development Goals (MDGs) to the Sustainable Development Goals (SDGs) [2]. The SDGs cover seventeen interconnected goals that define global quantifiable objectives across the social, economic and environmental dimensions of development. They aim to be a universal set of indicators and a reference framework to be leveraged by the global community to motivate policies and implementation by 2030.

This framework supports a long term transition towards more sustainable development. It fosters accountability while also promoting global collaboration. It is a tool to guide decision making but not in and of itself a prescriptive and actionable guide.

J. Hodson (✉)
AI for Good Foundation, El Cerrito, USA
e-mail: hodson@ai4good.org

A. Spezzatti
UC Berkeley, Berkeley, USA
e-mail: andy_spezzatti@berkeley.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_79

As a result, member states and institutions are free to pursue policies and programmes according to their own context, resources, and available *scientific evidence*.

Lack of data is one of the principal bottlenecks in making progress towards the Goals. This problem manifests itself both through (i) the difficulty of consistently measuring *progress*; and (ii) the difficulty of selecting appropriate interventions. The first problem is an active area of research, with web-based systems such as the *SDG Trendscanner*[1] recently maturing to the point of being generally useful. The second problem is not specific to the Sustainable Development Goals, it is an issue that plagues practically all empirical sciences, and has only become more evident as the web has grown.

Researchers do not do a good job of sharing data, which creates barriers to the reproduction of results, and makes it more difficult for scientists to build on each others' work. There are various motivating factors that might arguably explain the current equilibrium: (i) data has become the main innovation in many fields, so researchers seek to publish as many papers as possible on a data set before sharing it, (ii) many data sets are provided through special relationships with commercial entities who need to maintain control of their assets, (iii) departments, universities, states, and countries all may have competitive reasons to keep data private, and (iv) the available mechanisms for sharing data are often not completely open themselves, perhaps even being set up as commercial entities and having burdensome requirements or fees.[2]

This paper's main contribution is focussed on the issue of *data discovery* and *data sharing* to accelerate progress on the Sustainable Development Goals. Our aim is to provide a system that automatically identifies, aggregates, and describes data sets of relevance to each SDG, performs this task at a global scale, and provides ways to help researchers search for and obtain data sets for their work on Sustainable Development topics.

We leverage the intuition that most published research must provide sufficient details of the data sets used in order to pass any level of peer review. We do not expect authors of empirical work to do anything more than they already do to describe their data sets in their own papers, and publish their papers. It is the system's responsibility to read scientific publications, identify the data sets used, understand the data set in context, and decide whether or not it is relevant to a particular Sustainable Development Goal. Furthermore, the system must be able to associate details of a particular data set across multiple publications, and must be able to identify the owners and contact details of responsible parties where available. A limited number of data sets are truly open, in which case future researchers would be able to download them rather easily. More than 95% of the data sets identified by our system have no further

[1] The SDG Trendscanner, developed collaboratively by the United Nations Development Programme (UNDP), the RISE Research Institutes of Sweden, and the AI for Good Foundation, leverages Artificial Intelligence to aggregate as much global information as possible regarding progress towards each SDG and a variety of related trends. The system is available freely at https://sdg.trendscanner.online.

[2] See [9] for a more in-depth treatment of this topic. However, 'Open Data' services provided by academic publishers unfortunately often result in the public sharing of data tables from papers, rather than the underlying raw data used in analysis.

information available beyond what is written in the published work, despite being highly relevant to solving some of the world's most pressing challenges. In these cases, it is our hope that the data owners will be open to sharing their data when a peer scientist places a request relating to the Sustainable Development Goals, in return for appropriate citations.

The *SDG Data Catalogue* leverages Artificial Intelligence and Natural Language Processing techniques in order to "read" millions of academic papers indexed through the Open Academic Graph (OAG) project (see [8, 10]). The system retrieves full papers, identifies anchor mentions of data sets, and leverages sections of the document to identify further pertinent details, including the number of samples, ownership, and relevant attributes. Each paper is also classified as related or not to each of the Sustainable Development Goals, allowing data sets to be organised under each main category. Where the same data set is leveraged across multiple publications, the system is able to merge extracted information to retain consistent and cross-referenced records. Data set details can then be presented to users in a query-able and actionable format. We perform evaluation at each stage of the system's component pipeline, and find promising performance at each stage, with overall precision of the system's extractions nearing 89%, and recall being around 72% for data set discovery.

The remainder of this paper proceeds as follows: Sect. 2 describes the data set used, how it was obtained, and how it is cleaned/filtered in preparation for information extraction; Sect. 3 provides details of our annotation and information extraction pipelines; Sect. 4 describes the system architecture and usage; and Sect. 5 concludes.

2 Data

Our aim is to allow our algorithms to read as many papers as possible from the web and partner data sources. The more documents can be collected, analysed, and extracted, the more confident the system becomes in each individual data set extraction. The foundation for our approach is a flexible web-scraper that is both scalable, light-weight, and copyright-aware. Scalable such that we can keep up with the thousands of new documents available each day, light-weight so that we can deploy additional resources quickly, and copyright-aware so that we do not inadvertently retrieve materials subject to explicit policies that prohibit the use of web spiders or machine reading.

The basis for our web scraper is the extensive database of academic publications and references provided by the Open Academic Graph (OAG) project [8, 10]. The OAG data set contains information on approximately 240M unique global publications, spanning several decades, and across all domains. Each entry contains title, abstract, authors, citations, year of publication, and venue fields, among others.[3] Not

[3] A URL field is also often present, but tends to link to paywall content such as the IEEE and ACM digital libraries. We therefore do not use these links.

all fields contain valid data, but the coverage is very high (more than 95%) for the core entries we are interested in.

For each paper entry in OAG, we construct a search engine query containing the title and author last names. Query results are filtered to PDF files, and the top 5 results become part of a candidate list. We then proceed to attempting the download of each document in turn until we either (a) successfully retrieve a parse-able PDF file, or (b) reach the end of the candidate list. We leverage a proxy service[4] in order to parallelise our download queues up to 100 times and make maximal use of our servers' available bandwidth. With this approach, we retrieve valid PDF documents 65% of the time, and are able to download 10k valid documents per hour. For the purpose of this paper, we work with a subset of 5M valid PDF documents.

3 Parsing and Information Extraction

There is a large literature spanning several decades covering structured information extraction from textual documents. Previous work deals with (a) retrieving text in the correct intended order and format (e.g. [1]), (b) understanding the sections and relational structure of the document (e.g. [5]), and (c) extracting complex information such as tables, image captions, and citations (e.g. [3]).

Our present work falls somewhere across all of the above areas: we need to accurately extract the actual text content of the entire document; we then focus on the identification of the sections of a document that deal with our topic of interest (data sets); we employ specific entity recognition algorithms to extract complex relational information of interest, and aggregate this information across a large number of relevant documents.

Fortunately, for some of these areas there are existing implementations of sufficient quality that we can use off-the-shelf. For example, extracting text from PDF files is handled transparently for us by the Apache Tika Content Analysis and Extraction Toolkit.[5] We find that the validity of extracted text is sufficient (i.e. non-garbled and following the correct flow patterns of the original document) in 83% of cases. Therefore, it is important to note that our document processing pipeline is able to correctly process 54 documents out of every 100, before any advanced information extraction can take place (see Table 1).

[4] ScraperAPI Inc., scraperapi.com.

[5] https://tika.apache.org.

Table 1 Documents remaining in the pipeline at each stage of processing

	Documents in	Documents out	Proportion (%)
Web scraping	8,000,000	5,200,467	65
Text extraction	5,200,467	4,316,652	83
Data set identification	4,316,652	2,706,109	62

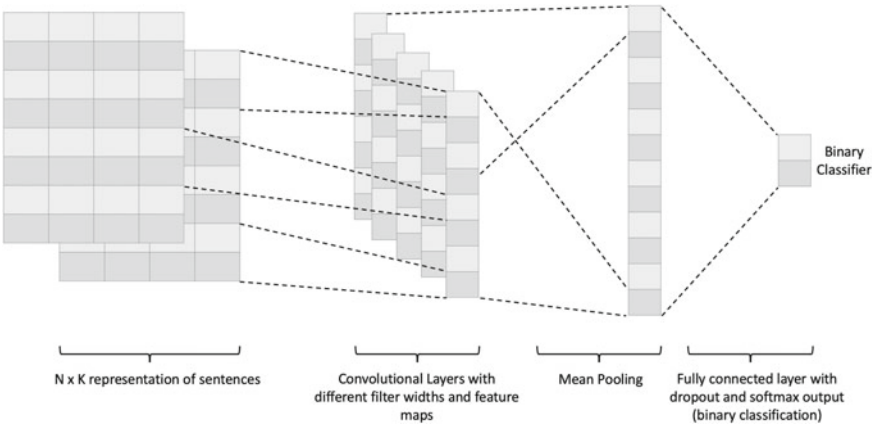


Fig. 1 Text classification neural architecture

3.1 Data Set Mentions

The first part of our information extraction pipeline isolates the segments of documents that contain some information about data sets. We take a boot-strapping approach to training a text classification model based on Convolutional Neural Networks implemented within the SpaCy NLP Library[6] as depicted in Fig. 1.

First, we extract all paragraphs containing the word 'data', in order to limit the number of candidates for supervised annotation. Next, we identify a series of linguistic patterns as seed terms for prioritising the discovery and annotation of positive examples of document chunks that are relevant to our objective.

Patterns of the form:

```
{the} {<ASCII>*} {data} {set}
{the} {<ASCII>*} {dataset}
...
```

which are highly likely to yield positive examples of data set mentions in the document text. We label these pattern-based examples manually, and use an initial set of

[6]See: Honnibal, M, Montani, I. SpaCy NLP Library, v2.3, June 2020. Online: https://github.com/explosion/spaCy.

Table 2 Identifying data set mentions: We are primarily interested in maximising recall at this stage

	Total documents	# data set mentions	Recall	Precision	F-1
Data Set Chunks CNN	10,000	5750	93.5%	48.6%	64%

500 annotations to train the text classification model over 10 epochs with a drop-out rate of 0.2. This yields a model with tenfold cross-validation F-1 score of 74.9%. We then leverage this model to suggest examples for annotation where the uncertainty is highest (where the model's confidence score among the two classes is lowest). We iteratively re-train our model using 300-dimensional word-embedding features trained on the CommonCrawl web-scale data set[7] with the GLobal VEctors for Word Representation (GloVe) procedure (see [7]). Table 2 shows the results when testing against a held-out data set of 10k documents manually annotated for mentions.

3.2 Named Entity Recognition

Now that we are able to reliably identify passages of text that contain details of data sets, we move to the next task: extracting specific identifying information about the data. We are interested in identifying the following fields when they appear:

- Data Set Name;
- Data Set Description;
- Owner;
- Number of samples;
- Attribute references.

In addition, we extract document metadata from Tika's Pipeline and OAG database which includes authors, institutions, year of publication, keywords, and publication venue. We also leverage a regular expression to match email addresses in the documents themselves in order to facilitate contacting those with access to or control over the data set.

Each entity of interest from the list above is manually identified through a concerted annotation exercise. We collect a minimum of 500 annotations for each entity within an active learning *model-in-the-loop* annotation pipeline (see Fig. 2).

Our Named Entity Recognition model is trained in the same way for each entity, leveraging a Stack-Long-Short-Term-Memory (sLSTM) approach following [6]. This transition-based approach is a relatively recent development in the field of NER, taking its inspiration from the literature on parsing. The results for each label

[7] https://commoncrawl.org.

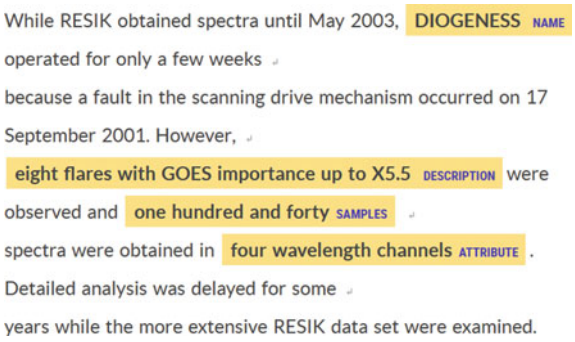


Fig. 2 An example NER annotation for several labels

Table 3 Performance of NER models, rounded to closest integer value

	Instances	Recall (%)	Precision (%)	F1 score (%)
Name	2725	72	89	80
Description	500	56	38	45
Owner	500	72	44	55
Samples	500	32	35	33
Attribute	742	12	24	16

type are presented in Table 3, and represent testing on a 20% held-out portion of each data set.

Performance varies greatly by label category, with worse performance observed for labels that are more heterogeneous. For instance, data set attributes vary widely in type and description, and mentions of data set sample size tend to be very similar to other numeric information that may appear nearby.

Fortunately, the system need not be too sensitive to one-shot extractions. Any data set that has been leveraged in multiple publications may build confidence in its knowledge of that data set through consistent extractions, which are merged.

4 A Sustainable Development Data Catalogue

While this paper represents experiments in building the pipeline for data set extraction to support a global data catalogue, it is worth dedicating some space to motivating and describing the ultimate goal of our work.

A pilot system that leverages text classification models from prior work [4] allows us to accurately associate each academic paper with zero or more of the Sustainable Development Goals. Being able to place each of the extracted data sets into a category, complete with the related papers, author and institution names (and more) provides

a data-set discovery system that has the potential to accelerate the application of Artificial Intelligence research on the Sustainable Development Goals. Of course, we wish to take this work further, building query and visualisation layers on top of the *SDG Data Catalogue* (as well as improving all aspects of the document processing pipeline), and allowing researchers to request access to and update details of data sets in the system with minimal overhead. In the end, we hop to create an easily navigable Wikipedia-like experience for keeping track of useful data.

5 Conclusion

The Coronavirus pandemics of 2019 and 2020 have highlighted how the global research community is able to quickly respond to crises and generate new knowledge to help identify solutions. However, the pandemics also highlighted how the creation of thousands of new papers each day on a single topic can cause more confusion than it might clarify.

Without appropriate intelligent systems it is not possible to keep up to date on new findings or know what data sets to use. The most visible Coronavirus-related data set was a compilation of research papers mentioning "coronavirus", followed closely by rudimentary aggregate statistics from the World Health Organisation (WHO). Clearly, we need to do better to provide our scientific community with the data and tools they need to support decision-making in challenging situations. The Sustainable Development Goals represent a large set of ever more challenging situations, and the *SDG Data Catalogue* is one more resource to help bridge the gap between a set of aspirational objectives, and a clear trajectory.

This paper describes a system that is under active development. The SDG Data Catalogue algorithmic infrastructure will continue collecting new research, improving its coverage and accuracy, and we will be adding new ways to aggregate, sort, and search the resulting information. Our ultimate objective is for the platform to offer incentives that shift the behaviour of research communities toward using their methods first and foremost for the benefit of humanity—bringing more visibility to those researchers and organisations that are most open, collaborative, and productive in developing new solutions to the SDGs.

References

1. Edmundson, H. P. (1969). New methods in automatic extracting. *Journal of the ACM (JACM)*, *16*(2), 264–285.
2. Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyamsundar, P., et al. (2013). Sustainable development goals for people and planet. *Nature, 495*, 305–307. https://doi.org/10.1038/495305a
3. Hodson, J. (2016). Table extraction from text documents. In C. Sammut & G. Webb (Eds.), *Encyclopedia of machine learning and data mining*. Boston, MA: Springer.

4. Hodson, J., & Novak, B. (2020). Trendscanning for the SDGs. AI for Good Foundation Report. https://ai4good.org/trendscanning. Accessed July 2, 2020.
5. Iwai, I., Doi, M., Yamaguchi, K., Fukui, M., & Takebayashi, Y. (1989). A document layout system using automatic document architecture extraction. In Bice, K., & Lewis, C. (Eds.), *CHI '89: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA (pp. 369–374). https://doi.org/10.1145/67449.67520
6. Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., & Dyer, C. (2016). Neural architectures for named entity recognition. *ArXiv Preprint*. arXiv:1603.01360.
7. Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. In A. Moschitti (Ed.), *Empirical methods in natural language processing '14*. NY, USA: Association for Computational Linguistics.
8. Sinha, A., Shen, Z., Song, Y., Ma, H., Eide, D., Hsu, B.-J., et al. (2015). An overview of Microsoft Academic Service (MAS) and applications. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15 Companion)* (pp. 243–246). New York, NY, USA: ACM.
9. Stuart, D., Baynes, G., Hrynaszkiewicz, I., Allin, K., Penny, D., Lucraft, M., et al. (2018). Practical challenges for researchers in data sharing. Springer Whitepaper. https://www.springernature.com/gp/open-research/open-data/practical-challenges-white-paper.
10. Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). ArnetMiner: Extraction and mining of academic social networks. *Proceedings of the Fourteenth ACM SIGKDD International Conference* (pp. 990–998).

Characterization of Heterogeneous Malware Contagions in Wireless Sensor Networks: A Case of Uniform Random Distribution

ChukwuNonso Nwokoye, Ikechukwu Umeh, and Obiajulu Ositanwosu

Abstract Most mathematical models representing infection spread in computer and wireless sensor networks (WSN) address only a specific kind of malware. However, researchers in the field of biosciences have modeled non-homogenous populations of contagions in a host. Motivated by this, we therefore propose the Vulnerable, Contagious due to virus, Contagious due to worm, Contagious due to trojan horse, Recovered with Inoculation ($e\text{-}VC_jRI$) epidemic model, to describe the propagation dynamics of multiple breeds of malware in a WSN. Beside distinctive infectiousness, the e-VCjRI model possesses expressions for communication range and distribution density, which are renowned WSN attributes that constituted the actual threshold parameter (R_o) alongside differential infectivity as result of worm, virus and trojan horse. Put another way, the study illustrated that the true R_o is the summation of each malware group's reproduction ratio. The Runge–Kutta order 4 and 5 numerical method was used to provide solutions for the system of differential equations, and afterward, the effects of the aforementioned WSN features were presented.

Keywords e-VCjRI model · Worm · Virus · Trojan horse · Communication range · Distribution density

C. Nwokoye (✉)
Nigerian Correctional Service, Awka, Nigeria
e-mail: chinonsonwokoye@gmail.com

I. Umeh · O. Ositanwosu
Nnamdi Azikiwe University, Awka, Nigeria
e-mail: ikumeh1@gmail.com

O. Ositanwosu
e-mail: oe.ositanwosu@unizik.edu.ng

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_80

1 Introduction

The emergence of cyberspace ushered in monumental rearrangements in this day and age, permitting prompt accessibility to information. Fundamentally, the Internet provides intrinsic support to the ever evolving cyberspace, wherein its continual operation offers an avalanche of prospects, ease and benefits. However, these advancements possess attendant threats of mischievous and hostile assaults from black hat hackers who aim to disrupt individuals' or institutions' daily usage of the Internet for meaning pursuits. Additionally, computer misapplication and nonconformity with regulations in workplaces also aid the malware intrusion into networks. Actually, these malicious codes with tremendously damaging tendencies terrorize the cyberspace's continuance and profitability [1]. Malicious attacks to communication networks arise from worms, viruses, rootkits and trojan horse. Behaviorally, self-replication is possible for worms, while viruses usually need human meddling or involvement, for instance opening an email attachment. In the case of trojan horses, they are developed, "to quietly monitor, to actively send out unencrypted sensitive information, or to actively destroy their host device" [2]. In the words of Nwokoye et al., [3] "WSN consists of sensor nodes which are distributed in a sensor field where they are connected to the sink, to track, record and send ambient territorial parameters to a data collector (or base station) through multihop infrastructureless transmission between neighboring sensor nodes." From the extant research in network security, sensors in a WSN are prone to attacks from worms, viruses and trojans, perhaps due to the fact that communication therein is done in an open manner. These malevolent assaults can be less invasive (i.e., breaching secrecy) or highly invasive (i.e., sensor functionality disruption or data integrity destruction). The ubiquitous nature of WSN attacks makes it excessively profitable to conceive successful approaches for dealing with the hazardous effects of malicious objects. Current discoveries show that these black hat hackers exploit the constrained resources (limited bandwidth, power, and range) of the sensors to unfurl infections throughout the network [3]. Consequently, to discern the dynamics and strategies of transmission, experts in information technology defense utilize epidemic models.

2 Related Works

Through epidemiological literature originating from public health, it was found that virtual malwares possess similar behaviors with agents of biological networks. Therefore, epidemic models with host compartmentalization based on health status have been utilized for comprehending infection propagation in all types of telecommunication networks. The addition of sensor vaccination to the obsolete SIR epidemic model to form our model was motivated by some remediation measures of WSN proposed by older epidemic models [4–6]. The challenge with these WSN models is that they characterized only one kind of malware. However, herein, we nurse the

assumption that multiple kinds of malware can exist in a sensor field—this concept is regarded as multi-group modeling. In the words of Driessche and Watmough, [7]. "'multigroup' usually refers to the division of a heterogeneous population into several homogeneous groups based on individual behavior…(where)…each group is then subdivided into epidemiological compartments". Probing through network epidemiological literature, we only found multi-group representations of computer networks and none for WSNs. For instance, Mishra and Singh [1] suggested the SIjRS model for control of malware in a computer network. Unlike the mass action incidence employed by Mishra and Singh [1], Mishra and Ansari [8] posited the electronic SIRS, with reinfection and differential vulnerability for self-replicating worms and intervention-based viruses. These multi-group computer network models are inappropriate for the true characterization of heterogeneous populations in a WSN because they are absent sensor features (i.e., transmission range and distribution density) described in these studies [9, 10] etc. Proposing a multi-group WSN model is essential because these works have illustrated the unarguable likelihood that more than one kind of malware can exist simultaneously in a network.

3 The e-VCjRI Model with Distinctive Infectiousness, Communication Range and Distribution Density

To describe the intricacies of spread for multiple malware infections in a WSN, we present the epidemic Vulnerable, Contagious due to virus, Contagious due to worm, Contagious due to trojan horse, Recovered with Inoculation (*e-VCjRI*) model, which is graphically illustrated as Fig. 1. The specific assumptions for our model go thus; Z is the addition of vulnerable sensors in the WSN, η is mortality rate as a result of reasons others than malware attack, σ is distribution density, r is communication range, β_1 is rate of virus infection, β_2 is rate of worm infection, β_3 is rate of trojan horse infection, e_1 is the death rate due to viral attack, e_2 is the death rate due to worm

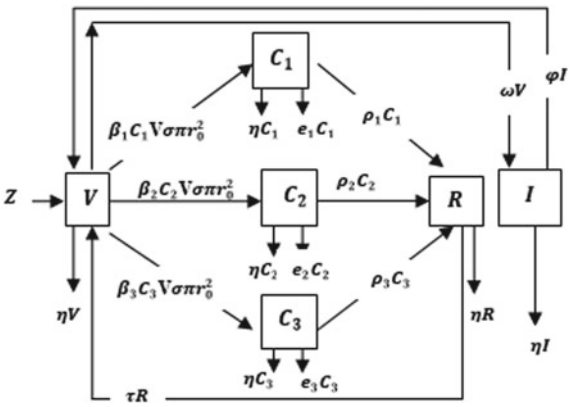


Fig. 1 Flow of three classes of WSN malware

attack, e_3 is the death rate due to trojan horse attack, ρ_1 is rate of recovery from viral infection, ρ_2 is rate of recovery from worm infection, ρ_3 is rate of recovery from trojan infection, τ is the rate at which nodes lose their transient immunity, φ is rate of transfer from the vaccination compartment to the vulnerable sensor compartment, and ω is the rate of vaccinating the vulnerable nodes against later infections. The total sensor population in the network at any time t is $N(t) = V(t) + C_1(t) + C_2(t) + C_3(t) + R(t) + I(t)$.

The e-VCjRI model for WSN is represented using the following system of differential equations;

$$\dot{V} = Z - \eta V - \omega V - \sum_{j=1}^{3} \beta_j C_j V \sigma \pi r_0^2 + \tau R + \varphi I$$

$$\dot{C} = \sum_{j=1}^{3} \beta_j C_j V \sigma \pi r_0^2 - (\eta + e_j + \rho_j) C_j; \quad j = 1, 2, 3. \tag{1}$$

$$\dot{R} = \sum_{j=1}^{3} \rho_j C_j - (\eta + \tau) R$$

$$\dot{I} = \omega V - (\eta + \varphi) I$$

3.1 Existence of Equilibrium

By equating the system of Eqs. (1) to zero; $\dot{V} = 0$; $\dot{C}_1 = 0$; $\dot{C}_2 = 0$; $\dot{C}_3 = 0$; $\dot{R} = 0$; $\dot{I} = 0$, two equilibriums emerge, namely contagion-free equilibrium (CFE) and the endemic equilibrium (EE). Through straightforward computation, we obtain the following solutions for the individual compartments at the CFE;

$$V^0 = \frac{Z(\eta + \varphi)}{\eta(\eta + \varphi + \omega)}, C_1^0 = 0, C_2^0 = 0, C_3^0 = 0, R^0 = 0, I^0 = \frac{Z\omega}{\eta(\eta + \varphi + \omega)}$$

While the EE is as follows, where $\sigma \pi r_0^2$ is substituted as J;

$$V_j^* = \sum_{j=1}^{3} \frac{\eta + e_j + p_j}{\beta_j J}$$

$$C_j^* = \sum_{j=1}^{3} \frac{(\eta + \tau)\big(-\eta(\eta + \varphi + \omega)\big(\eta + e_j + p_j\big) + Z(\eta + \varphi)\beta_j J\big)}{(\eta + \varphi)\big((\eta + \tau)\big(\eta + e_j\big) + \eta p_j\big)\beta_j J}$$

$$R_j^* = \sum_{j=1}^{3} \frac{p_1\left(-\eta(\eta+\varphi+\omega)\left(\eta+e_j+p_j\right)+Z(\eta+\varphi)\beta_j J\right)}{(\eta+\varphi)\left((\eta+\tau)\left(\eta+e_j\right)+\eta p_j\right)\beta_j J}$$

$$I_j^* = \sum_{j=1}^{3} \frac{\omega\left(\eta+e_j+p_j\right)}{(\eta+\varphi)\beta_j J}$$

3.2 The Threshold Parameter

This quantity is most times called the reproduction ratio, i.e., the anticipated amount of newly infected cases from a characteristically infectious host in an entirely vulnerable population [11]. The threshold quantity for the e-VCjRI model is

$$R_o = \sum_{j=1}^{3} \frac{\beta_j J}{\eta+e_j+p_j}$$

Actually, the threshold quantity presented above involves the rates of virus, worm and Trojan infections, the rate of deaths as a result of these malwares and other reasons aside malicious code attack (i.e., software or hardware failure) as well as the communication range and distribution density. This shows that the true R_0 in a multi-group context is the summation of each malware group's reproduction ratio.

4 Numerical Simulation and Discussion

Modeling heterogonous infectious populations was viewed as an initial value problem, and therefore, to provide solutions for the proposed system (1), the Runge–Kutta–Fehlberg method was used. It is assumed that the WSN has the following initial values $V = 100$; $C_1 = 30$; $C_2 = 20$; $C_3 = 35$; $R = 0$; $V = 0$. The other values used the simulation experiments include $\eta = 0.006$, $Z = 0.095$, $\beta_1 = 0.07$, $\beta_2 = 0.05$, $\beta_3 = 0.039$, $\tau = 0.006$, $e_1, e_2, e_3 = 0.889$, $\rho_1 = 0.009$, $\rho_2 = 0.008$, $\rho_3 = 0.007$, $\omega = 0.150$, $\varphi = 0.06$, $\sigma = 0.5$ and $r = 1$. Interestingly, increasing the range to 2 automatically increased the infection in the WSN, and Fig. 2 portrays the behavior of the compartments. However, for Fig. 3, the recovery rates and vaccination rate were increased to $\rho_1 = 0.09$, $\rho_2 = 0.08$, $\rho_3 = 0.07$ and $\omega = 1.150$, respectively. The visibility of this R and I compartments (at Figs. 3, 4 and 6) shows that network managers should proportionally heighten these countermeasures whenever range is varied. Figure 4 depicts the impact of transmission range on the subgroups of the contagious class, while the phase plot of Fig. 5 depicts the impact of several communication ranges on node vulnerability, recovery and inoculation. While performing the simulation,

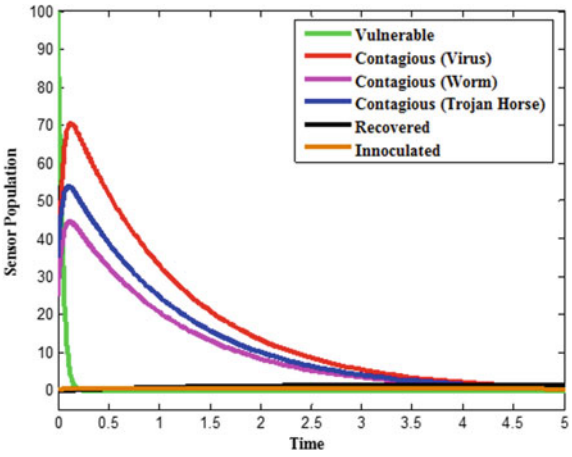


Fig. 2 Time histories for all compartments (when recovery and inoculation rates are kept at $\rho_1 = 0.009$, $\rho_2 = 0.008$, $\rho_3 = 0.007$ and $\omega = 0.150$, respectively)

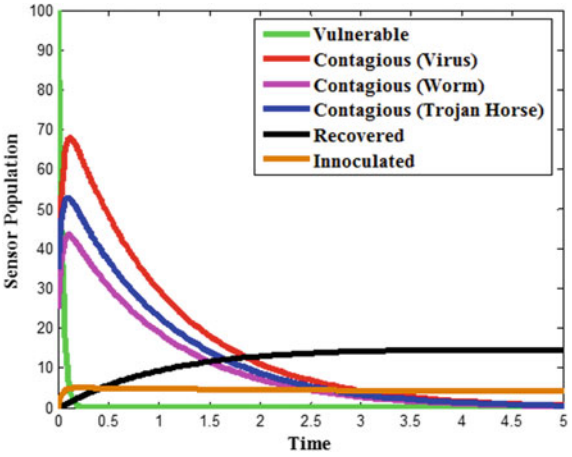


Fig. 3 Time histories at $\rho_1 = 0.09$, $\rho_2 = 0.08$, $\rho_3 = 0.07$ (R) and $\omega = 1.150$ (I))

density was kept at 0.5, while the range assumed these values 2, 7 and 12. Keeping the range constant, we portray the effect of density on the sub-compartments of the contagious class (Fig. 6). More so, we also presented as Fig. 7 its effect on the vulnerability, recovery and inoculation. Note that Figs. 6 and 7 showed the behavior of the WSN by varying density at 0.5, 1.0 and 1.5, i.e., the implication is that they display some sensitivity to the increase of sensor density for all the compartments involved. Consequently, this implies that with increasing infection, sensor recovery and inoculation is greatly reduced.

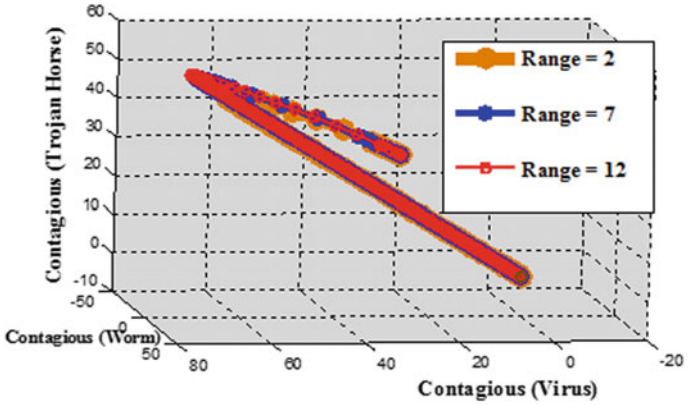


Fig. 4 Dynamical behavior of the contagious class showing the impact of communication range

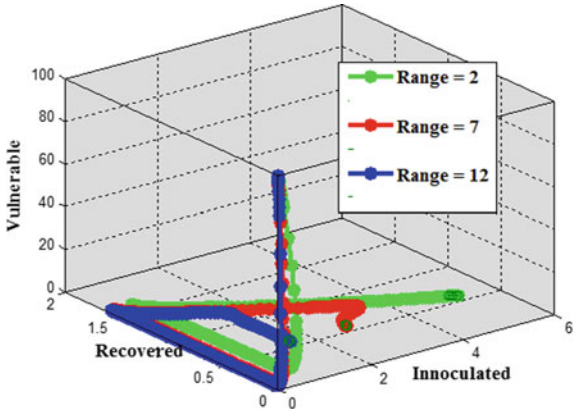


Fig. 5 Impact of communication range on sensor vulnerability, recovery and inoculation

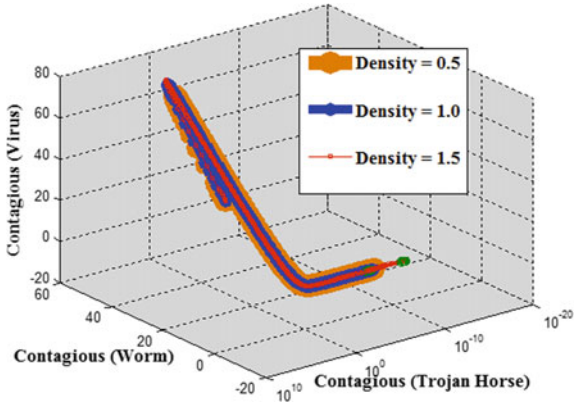


Fig. 6 Effect of density on the contagious sub-compartments for virus, worm and trojan horse

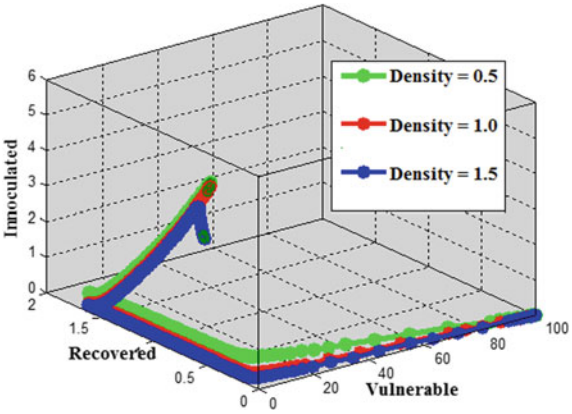


Fig. 7 Effect of distribution densities on vulnerable, recovered and inoculated sensor nodes

5 Conclusion

Our work was inspired by multi-group representations of biological diseases in the field of mathematical biosciences and the heavy tendency that viruses, worms and trojan horses might exist in a network. Therefore, we posited the e-VCjRI epidemic model. This is very useful in WSN since most differential epidemic models characterize one form of malware contagion. Furthermore, we derived the solutions at existent equilibria as well as the threshold parameter, which is a summation of individual reproduction ratio of sub-compartments. Simulation experiments were performed using the suitable numerical method. Specifically, while keeping the density constant at 0.5, the range was varied using the following values; 2, 7 and 12. On the other hand, we kept the range constant at 1 and then changed density in this manner 0.5, 1.0 and 1.5. The 3D phase plots allowed the perfect portrayal of the effects on the compartments. The insights gained from this study would positively impact decisions of network managers. In the future, we would explore the impact of node exposure and sensor isolation for subsequent remediation, of course, while considering the multi-group idea.

References

1. Mishra, B. K., & Singh, A. K. (2012). SIjRS e-epidemic model with multiple groups of infection in computer network. *International Journal of Nonlinear Science, 13,* 357–362.
2. Jalalitabar, M. (2014). Trojans in wireless sensor networks. Georgia State University, US. https://scholarworks.gsu.edu/cs_theses/78.
3. Nwokoye, C. H., Mbeledogu, N., Umeh, I. I., & Ejimofor, I. A. (2017). Modeling the effect of network access control and sensor random distribution on worm propagation. *International Journal of Modern Education and Computer Science, 9,* 49–57.
4. Nwokoye, C., & Umeh, I. (2018). Analytic-agent cyber dynamical systems analysis and design method for modeling spatio-temporal factors of malware propagation in wireless sensor

networks. *MethodsX, 5,* 1373–1398.
5. Zhang, Z., Kundu, S., & Wei, R. A. (2019). Delayed epidemic model for propagation of malicious codes in wireless sensor network. *Mathematics, 7,* 1–17.
6. Batista, F. K., Martín del Rey, A., & Queiruga-Dios, A. (2020). A new individual-based model to simulate malware propagation in wireless sensor networks. *Mathematics., 8,* 1–23.
7. Driessche, P., & Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences, 180,* 29–48.
8. Mishra, B. K., & Ansari, G. M. (2012). Differential epidemic model of virus and worms in computer network. *International Journal of Network Security, 14,* 149–155.
9. Feng, L., Song, L., Zhao, Q., Wang, H. (2015) Modeling and stability analysis of worm propagation in wireless sensor network. *Mathematical Problems in Engineering*, 1–8.
10. Wang, Y., & Yang, X. (2013). Virus spreading in wireless sensor networks with a medium access control mechanism. *Chinese Physics B, 22,* 40200–40206.
11. Diekmann, O., Heesterbeek, J. A., & Metz, J. A. (1990). On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology, 28,* 365–382.

Retraction Note to: A Computer Vision-Based Approach for Subspace Clustering and Lagrange Multiplier Optimization in High-Dimensional Data

K. R. Radhika, C. N. Pushpa, J. Thriveni, and K. R. Venugopal

Retraction Note to: Chapter "A Computer Vision-Based Approach for Subspace Clustering and Lagrange Multiplier Optimization in High-Dimensional Data" in: S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notesin Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_14

The publisher would like to alert readers that the conference paper [1] was retracted because of a production error resulting in duplicate publication in the same book series in different volumes. The correct citation for this article should be from the original publication [2].

[1] Radhika K.R., Pushpa C.N., Thriveni J., Venugopal K.R. (2021) A Computer Vision-Based Approach for Subspace Clustering and Lagrange Multiplier Optimization in High-Dimensional Data. In: Fong S., Dey N., Joshi A. (eds) ICT Analysis and Applications. Lecture Notes in Networks and Systems, vol 154. Springer, Singapore. https://doi.org/10.1007/978-981-15-8354-4_14

[2] Radhika K.R., Pushpa C.N., Thriveni J., Venugopal K.R. (2020) A Computer Vision Based Approach for Subspace Clustering and Lagrange Multiplier Optimization in High-Dimensional Data. In: Fong S., Dey N., Joshi A. (eds) ICT Analysis and Applications. Lecture Notes in Networks and Systems, vol 93. Springer, Singapore. https://doi.org/10.1007/978-981-15-0630-7_43

The retracted version of this chapter can be found at https://doi.org/10.1007/978-981-15-8354-4_14

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4_81

Author Index

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 154, https://doi.org/10.1007/978-981-15-8354-4